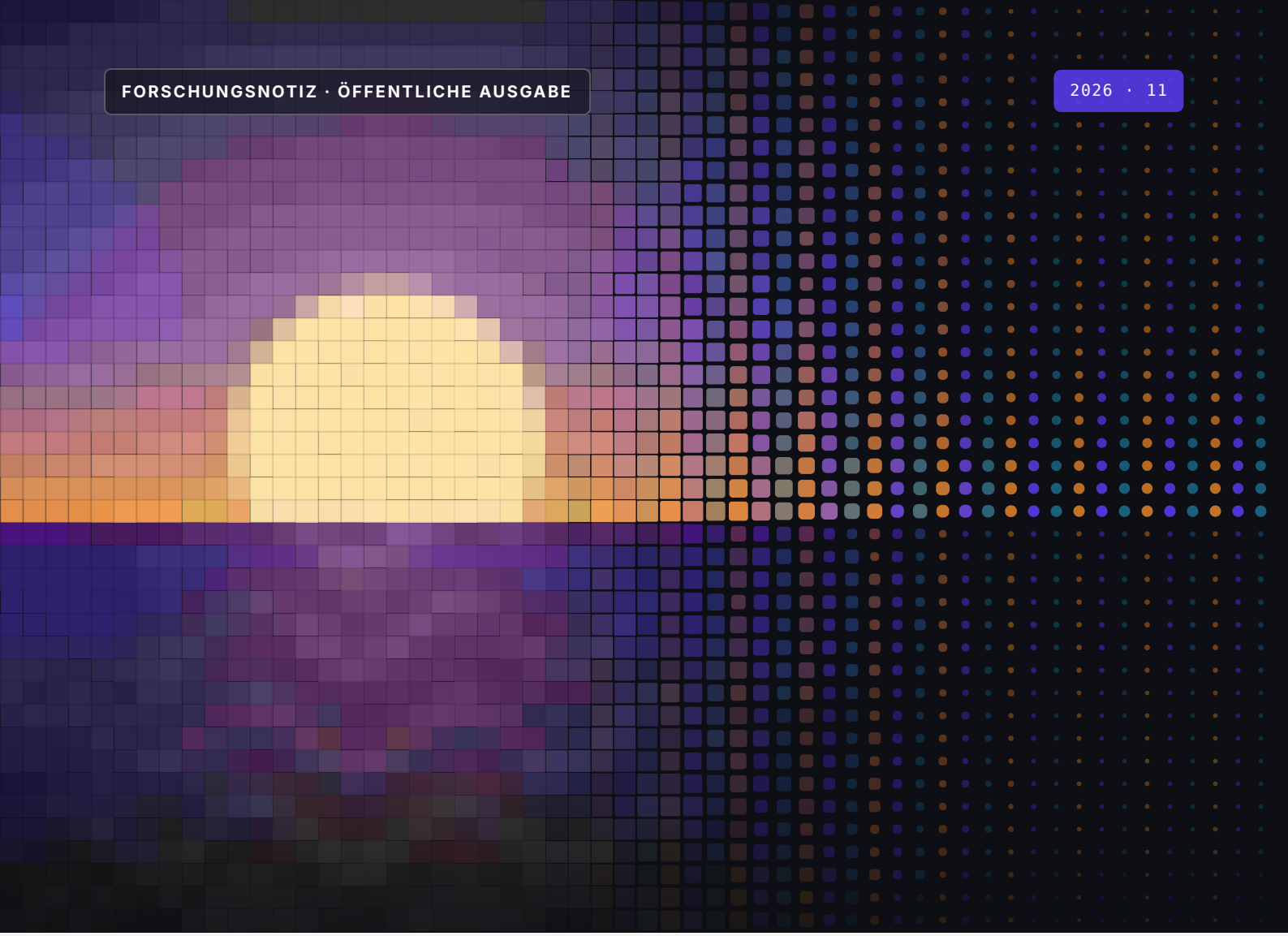


FORSCHUNGSNOTIZ · ÖFFENTLICHE AUSGABE

2026 · 11

GELERNTE KOMPRESSION · WIEDERGABE × SILIZIUM

Neuronale Codierung 2026

Im Einsatz rund um den Codec, nicht an seiner Stelle

Wiedergabe · Fotografie · Telefonie · Generative KI · Spiele · Silizium

DAS FAZIT

Neuronale Netze arbeiten bereits an beiden Enden der Codecs, mit denen die Welt ihre Videos überträgt. Neuronale End-to-End-Codecs gibt es nur in geschlossenen Systemen, und die Bildschirme des Massenmarkts warten auf H.267.

Tan Shuai

Technologe und Gründer · KI-Forscher und -Entwickler

tanshuai.com · i@tanshuai.com

Datenstand: 12. November 2026

Beträge in US-Dollar

INHALT

Inhalt



◆ **Zusammenfassung** — Bereits im Mainstream, aber rund um den Codec — 04

01 **Abgrenzung und Methode** — Was „neuronale Codierung“ hier bedeutet und wie ich vorgegangen bin — 07

02 **Wiedergabe und Streaming** — Wiedergabe und Streaming: neuronal an beiden Enden, Standard in der Mitte — 09

03 **Fotografie und Bilder** — Fotografie und Bilder: neuronal in der Kamera, klassisch im Web — 12

04 **Telefonie und Konferenzen** — Telefonie und Konferenzen: wo neuronale End-to-End-Codecs tatsächlich im Einsatz sind — 13

05 **Generative KI** — Generative KI: der größte Einsatz, den niemand Codec nennt — 15

06 **Spiele und Grafik** — Spiele und Grafik: neuronale Rekonstruktion ist etabliert, neuronale Assets stehen am Anfang — 16

07 **Spezialverbindungen** — Spezialverbindungen: Satelliten, Drohnen und Roboter — 18

08 **Silizium** — Der Blick aufs Silizium: Energie, Rechenaufwand, Speicher und Exaktheit — 19

09 **Standards und Patente** — Standards und Patente: wo neuronale Werkzeuge lizenzierbar werden — 24

10 **Wer verdient** — Wer verdient, und wo ein Unternehmen ansetzen kann — 25

11 **Behauptungen und Belege** — Behauptungen, die Sie hören werden, und was die Belege sagen — 27

12 **Wegmarken** — Worauf zu achten ist, 2027 bis 2030 — 28

13 **Offene Fragen** — 29

◆ **Über den Autor** — Tan Shuai — 30

◆ **Quellen** — 31

ABBILDUNGEN

Abbildungen



ABB. 01	Wo neuronale Codierung 2026 im Einsatz ist	06
ABB. 02	Drei Wege neuronaler Netze in den Codec	08
ABB. 03	Vom fertigen Standard bis zum Massenmarkt vergeht der Großteil eines Jahrzehnts	10
ABB. 04	Bei Anrufen schlägt ein neuronaler Codec Hardware-HEVC um 65–76 %	14
ABB. 05	Mit dem Stack aus einer Hand verbreitet sich neuronale Rekonstruktion schnell	17
ABB. 06	Energie pro decodiertem Pixel: Festverdrahtete Decoder liegen eine Größenordnung vorn	20
ABB. 07	Ein neuronales Werkzeug kostet Zehntausende MACs pro Pixel, ein ganzer neuronaler Decoder Hunderttausende	21
ABB. 08	Die Aufteilung in Serienchips: festverdrahtete Hardware für den Bitstrom, eine neuronale Einheit für die Pixel	22
ABB. 09	Wer 2026 an neuronaler Codierung verdient	25

Kennzeichnungen im Text

- **Schätzung** Rechnung oder Schlussfolgerung des Autors aus zitierten Daten
- **Anbieterangabe** vom Unternehmen veröffentlicht, das das Produkt verkauft, nicht unabhängig überprüft
- **Analyse** Interpretation der Belege durch den Autor

Dies ist eine Übersetzung der englischen Ausgabe; bei Abweichungen gilt die englische Fassung.

Datenstand: 12. November 2026 · Beträge in US-Dollar

ZUSAMMENFASSUNG

Bereits im Mainstream, aber rund um den Codec

DAS FAZIT

Neuronale Codierung ist 2026 bereits im Mainstream angekommen, aber sie arbeitet rund um die Standard-Codecs, nicht an ihrer Stelle.

≈75 %
des mobilen Datenverkehrs ist Video, Ende 2025 (Ericsson)

1,26 Mrd.
2025 ausgelieferte Smartphones, jedes davon decodiert Video in festverdrahteter Hardware

≈12×
Energie pro Pixel gegenüber einem gemessenen HEVC-Decoderchip mit DRAM: das einzige veröffentlichte neuronale Decoder-Chipdesign

≈22 TOPS
Dauerrechenleistung für einen neuronalen Decoder der DCVC-RT-Klasse bei 1080p30 Schätzung

75,5 %
weniger Bitrate als Hardware-HEVC braucht MLVC bei 360p in Anrufen, auf Microsofts eigenem Testdatensatz
Anbieterangabe

Ende 2029
H.267 Version 1, der erste Standard, der neuronale Werkzeuge wahrscheinlich auf jeden Bildschirm bringt

Wo sie im Einsatz ist

Neuronale Netze laufen heute an beiden Enden der Codecs, die fast alle Videos der Welt transportieren. Am Eingang formen sie das Bild: KI-Bildsignalprozessoren in Smartphone-Kameras und ein neuronaler Downscaler in der Encoding-Pipeline von Netflix. Am Ausgang reparieren sie es: YouTubes Super Resolution auf Fernsehern, Super-Resolution bei Meta und TV-Prozessoren, die mit Hunderten neuronaler Netze werben. Dazwischen ist der Bitstrom weiterhin AV1, HEVC, AVC oder VVC und wird von festverdrahteter Hardware decodiert. Kein Streamingdienst schickt ein neuronales Videoformat an Fernseher oder Smartphones; wo heute ein neuronaler Videodecoder auf einem Gerät läuft, läuft er in der eigenen App eines einzelnen Unternehmens.

Geschlossene Systeme

Neuronale End-to-End-Codecs, bei denen Encoder und Decoder gelernt sind, gibt es nur in geschlossenen Systemen, in denen ein Unternehmen beide Enden kontrolliert und auf einen Standard-Codec ausweichen kann. Microsoft betreibt seit 2021 einen KI-Audiocodec in allen Zweiergesprächen in Teams und Skype und hat im Juli 2026 MLVC als Open Source veröffentlicht, einen neuronalen Videocodec, der derzeit in Peer-to-Peer-Anrufen in Teams eingeführt wird: Bei 360p braucht er für die gleiche bewertete Qualität etwa 75 % weniger Bitrate als Hardware-HEVC. Jedes Modell zur Sprach- und Videogenerierung läuft auf einem gelernten Codec, von Sprach-Tokenizern mit 12,5 Hz bis zu Video-VAEs mit einem Kompressionsverhältnis von 1:192. Neuronale Texturkompression wurde in einem AAA-Spiel ausgeliefert. Das sind echte Einsätze, und sie teilen eine Bedingung: Niemand sonst muss den Stream decodieren.

Zeitplan

Die Wiedergabe neuronaler Bitströme im Massenmarkt wartet auf einen Standard und einen Chipzyklus. Der erste Standard, der wahrscheinlich neuronale Werkzeuge enthalten wird, der Nachfolger von VVC, den ich H.267 nenne, hat im Juli 2026 seinen Call for Proposals veröffentlicht, wertet die Einreichungen im Januar 2027 aus und zielt auf eine erste Version Ende 2029. AV1 brauchte von der fertigen Spezifikation bis zu 30 % des Netflix-Streamings fast acht Jahre. Nach diesem Muster kommen die ersten H.267-Chips um 2031 und der breite Einsatz Mitte der 2030er-Jahre Schätzung. Der Codec, der als Nächstes die Bildschirme erreicht, ist derweil AV2, im Mai 2026 fertiggestellt und ohne neuronalen Decoder, und drei IP-Anbieter haben binnen drei Wochen Kunden für AV2-Decoder bekanntgegeben.

Silizium

Der Grund liegt im Silizium. Das einzige veröffentlichte Chipdesign eines neuronalen Videodecoders braucht 14,6 nJ pro Pixel, etwa das 12-Fache eines gemessenen HEVC-Decoderchips samt DRAM und etwa das 47-Fache von dessen Kern allein. Microsofts Echtzeit-Forschungscodec DCVC-RT braucht nach meiner Zählung anhand seiner eigenen Tabelle pro Modul etwa 175 Tausend Multiply-Accumulate-Operationen pro decodiertem Pixel, also rund 22 TOPS Dauerrechenleistung für 1080p30 Schätzung: etwa die Hälfte der 40 TOPS, die einen KI-PC definieren, verbraucht für einen einzigen Videostream. Ein neuronaler Decoder muss zudem auf jedem Gerät bitidentische Ergebnisse liefern, sonst bricht seine Entropiedecodierung zusammen. Deshalb opfert MLVC Kompression, um auf allen NPUs identisch zu decodieren, und läuft mit 540p, nicht mit 4K.

Geschäft

Geschäftlich folgt das Geld den Standards und den Plattformen, nicht eigenständigen neuronalen Codecs. Plattformen behalten die Einsparungen aus neuronaler Vor- und Nachverarbeitung. IP-Anbieter lizenzieren festverdrahtete Decoder für jeden neuen Standard, dazu NPUs für die Bildverarbeitung. Patentpools kassieren pro Gerät: 0,50 US-Dollar pro Smartphone und 1,00–1,50 US-Dollar pro 4K-TV für VVC in den USA und anderen Märkten der Region 1. Die drei bekanntesten Start-ups für neuronale Codecs, WaveOne, iSIZE und Deep Render, wurden von Apple, Sony und InterDigital gekauft. Die offenen Chancen liegen in neuronaler Vor- und Nachverarbeitung, in Produkten für geschlossene Systeme, in Silizium, das einen festverdrahteten Codec mit einer NPU kombiniert, und in der Evaluierungs- und Standardisierungsarbeit dazwischen.

Frage	Einschätzung	Kennzahlen (Quellen in den Kapiteln)
① **Wo ist neuronale Codierung heute breit im Einsatz?**	Rund um den Codec: Kamera-ISPs, Downscaling vor dem Encoder, Super-Resolution bei der Wiedergabe. End-to-End nur in geschlossenen Systemen: Telefonie, Spiel-Assets, Tokenizer generativer KI	Video ≈75 % der mobilen Daten; 1,26 Milliarden Smartphones und ≈200 Millionen Fernseher 2025 ausgeliefert, alle mit Decodierung in festverdrahteter Hardware; DLSS 4 in über 250 Spielen und Apps
② **Werden neuronale Codecs AV1, HEVC und VVC bei der Wiedergabe ersetzen? Wann?**	Nicht vor einem Standard und einem Gerätezyklus. H.267 ist der erste realistische Träger und wird wahrscheinlich klassische und neuronale Werkzeuge mischen	Call for Proposals für H.267 im Juli 2026, Auswertung im Januar 2027, Version 1 Ende 2029; breiter Einsatz Mitte der 2030er-Jahre Schätzung
③ **Was kostet das in Silizium?**	Eine Größenordnung mehr Energie pro Pixel als heutige festverdrahtete Decodierung, dazu Speicherverkehr und die Anforderung der Bitgenauigkeit	14,6 nJ/Pixel (neuronal) gegenüber 1,19 (HEVC, mit DRAM); ≈175 kMAC pro decodiertem Pixel ≈ 22 TOPS bei 1080p30 Schätzung

Frage	Einschätzung	Kennzahlen (Quellen in den Kapiteln)
④ **Was erlauben Standards und Patente?**	Neuronale Postfilter lassen sich in VVC-Streams bereits signalisieren; JPEG AI ist ein ISO-Standard; H.267 entscheidet, welche neuronalen Werkzeuge lizenzierbar werden	ECM (klassisch) ≈28 % gegenüber VVC; neuronale Loop-Filter 6–14 % bei 10–22+ kMAC/Pixel; ByteDance mit 645 Patentfamilien in der Arbeit jenseits von VVC
⑤ **Wer verdient, und wo kann ein Unternehmen ansetzen?**	Plattformen, IP-Anbieter, Patentinhaber. Reine Start-ups für neuronale Codecs wurden übernommen. Ansetzen bei Vor- und Nachverarbeitung, geschlossenen Systemen, Silizium aus Codec plus NPU, Evaluierung	VVC-Pool 0,50 US-Dollar pro Smartphone; AV2-Decoder-IP wenige Wochen nach der Spezifikation lizenziert; drei Übernahmen, Bedingungen nicht veröffentlicht
⑥ **Welche Behauptungen verdienen Skepsis?**	Gewinne, die gegenüber alten Codecs, bei niedriger Auflösung, ohne Decodieraufwand oder aus Demos statt aus ausgelieferten Produkten angegeben werden	MLVC: 88 % weniger als H.264, aber 65–76 % weniger als Hardware-HEVC; Texturen: 85 % VRAM-Ersparnis in einer Demo, etwa 30 % im ausgelieferten Spiel

ABB. 01

Wo neuronale Codierung 2026 im Einsatz ist

Neuronale Netze laufen bereits an beiden Enden der gängigen Codecs. Neuronale End-to-End-Codecs gibt es nur dort, wo ein Unternehmen Encoder und Decoder kontrolliert.

	Rund um einen Standard-Codec Vorverarbeitung, Bildverbesserung, Upscaling	Neuronaler End-to-End-Codec gelernter Encoder und Decoder	Im Standard neuronale Werkzeuge in Standard-Bitstrom oder API
01 **Wiedergabe und Streaming**	BREIT IM EINSATZ Netflix-Downscaler; Super-Resolution bei YouTube und Meta; KI-Upscaler in TVs	NICHT IM EINSATZ Kein Streamingdienst sendet ein neuronales Videoformat; TVs und Smartphones decodieren Standard-Codecs in Hardware	IN STANDARDISIERUNG H.267 Version 1 kommt Ende 2029; neuronale Werkzeuge gehören zu den Kandidaten
02 **Fotografie und Bilder**	BREIT IM EINSATZ KI-ISPs in Flaggschiff-Smartphones; voll neuronaler ISP als IP, gezeigt auf der CES 2026	DEMOS, PILOTPROJEKTE JPEG AI ist ISO-Standard: Smartphone-Demos, kein Produkt	NICHT IM EINSATZ Das Format, das an Boden gewinnt, ist klassisch: JPEG XL in Chrome 155 standardmäßig aktiv
03 **Telefonie und Konferenzen**	BREIT IM EINSATZ Empfängerseitige Super-Resolution und neuronale Geräuschunterdrückung	BREIT IM EINSATZ Neuronales Audio in Teams-Zweiergesprächen seit 2021; MLVC-Video wird in Teams eingeführt	BEGRENZT IM EINSATZ Opus 1.5 ergänzt optionale, abwärtskompatible ML-Werkzeuge
04 **Generative KI**	—	BREIT IM EINSATZ Jedes Video- und Voice-Modell nutzt einen gelernten Tokenizer, z. B. einen Video-VAE mit 1:192	—
05 **Spiele und Grafik**	BREIT IM EINSATZ Neuronales Upscaling und Frame Generation: allein DLSS 4 in über 250 Spielen und Apps	BEGRENZT IM EINSATZ Neuronale Texturen in einem AAA-Spiel ausgeliefert; NVIDIAs SDK ist in der Beta	IN STANDARDISIERUNG DirectX erhält Matrixoperationen für Shader (Vorschau)
06 **Spezialverbindungen: Satellit, Roboter**	DEMOS, PILOTPROJEKTE Serverseitige Restaurierung von Feeds mit niedriger Bitrate	DEMOS, PILOTPROJEKTE Verlusttolerante neuronale Codecs in Forschung und Pilotprojekten	—

Quelle: Im Bericht zitierte Dokumente von Unternehmen, Standardisierungsgremien und Entwicklern; Stand 12. November 2026. Die Zeilen sind nach der Reichweite jedes Bereichs geordnet.

ABGRENZUNG UND METHODE

01 Was „neuronale Codierung“ hier bedeutet und wie ich vorgegangen bin

Drei Wege neuronaler Netze in den Codec

„Neuronale Codierung“ umfasst drei verschiedene Dinge, und die meisten Debatten darüber entstehen, weil man sie vermischt.

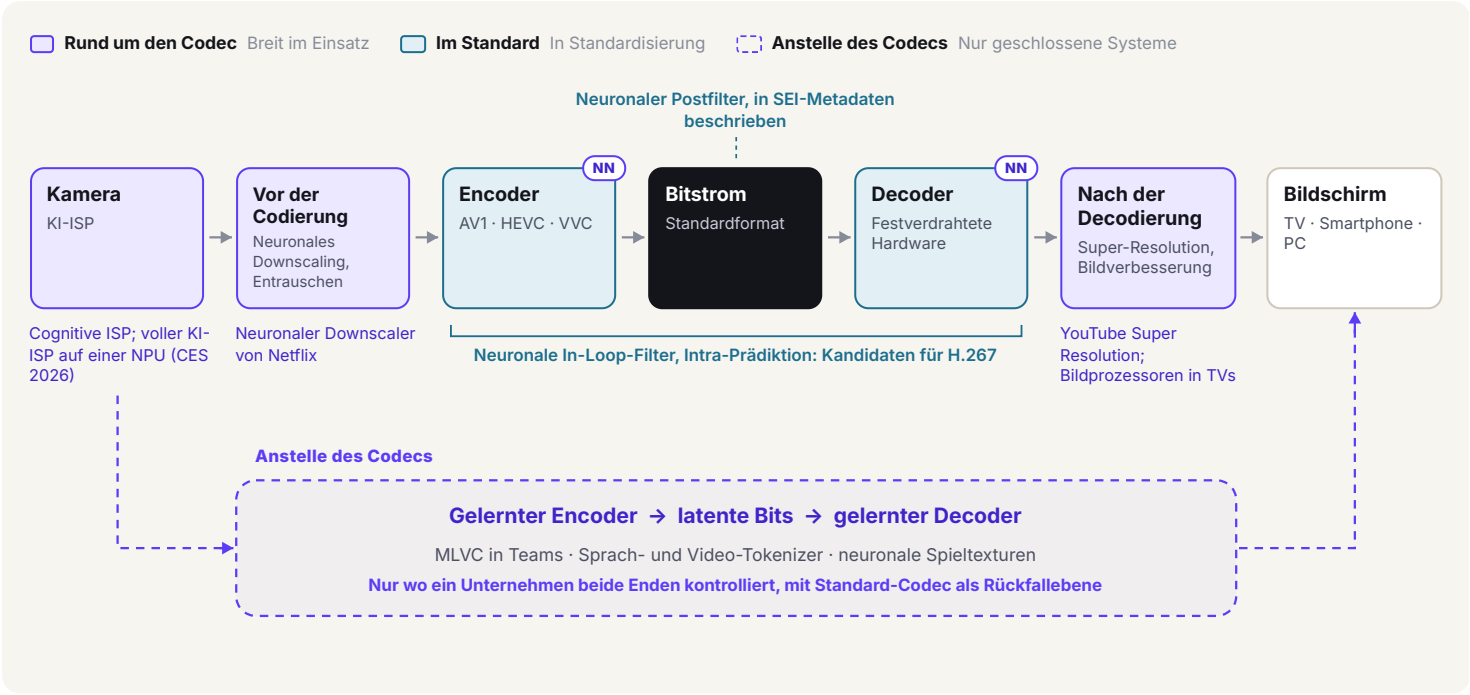
01 **Rund um den Codec.** Ein Netz verarbeitet Pixel vor der Codierung oder nach der Decodierung, und der Bitstrom bleibt standardkonform. Beispiele: ein KI-Bildsignalprozessor (ISP) in einer Kamera, der neuronale Downscaler, den Netflix seinem Encoder vorschaltet, Super-Resolution in einem Fernseher. Jedes Gerät kann den Stream abspielen; das Netz verändert nur, was hineingeht oder was herauskommt.

02 **Im Standard.** Ein Netz wird zum Werkzeug in einem Standard-Codec, etwa als neuronaler In-Loop-Filter oder Intra-Prädiktor, den jeder Decoder ausführen muss; es muss deshalb exakt spezifiziert werden, typischerweise in Ganzzahlarithmetik. Eine leichtere Form ist ein neuronaler Postfilter, den der Bitstrom in Metadaten beschreibt und den Decoder wahlweise anwenden können.

03 **Anstelle des Codecs.** Encoder und Decoder werden beide End-to-End gelernt. Der Bitstrom ist eine Menge quantisierter latenter Werte, die nur das passende Netz decodieren kann. Beispiele: Microsofts MLVC in Teams, Sprach-Tokenizer in Voice-Modellen, neuronale Texturen in einem Spiel.

ABB. 02 **Drei Wege neuronaler Netze in den Codec**

Das meiste, was 2026 im Einsatz ist, arbeitet rund um einen Standard-Codec. Neuronale Werkzeuge im Standard werden gerade spezifiziert. Neuronale End-to-End-Codecs ersetzen den Codec nur dort, wo ein Unternehmen beide Enden kontrolliert.



Quelle: Synthese des Autors aus den Quellen dieses Berichts; Stand 12. November 2026.

Die erste Art braucht niemandes Erlaubnis. Die zweite braucht einen Standard. Die dritte braucht die Kontrolle über beide Enden. In dieser Reihenfolge erreichen sie die Menschen, und das erklärt das meiste, was folgt.

Geordnet nach der realen Welt, nicht nach der Forschungsagenda

Forschungsarbeiten untersuchen meist die dritte Art, weil dort die Kompressionsrekorde liegen. Dieser Bericht ordnet die Bereiche danach, wie viele Menschen sie erreichen, weil das entscheidet, wohin Geld und Entwicklungsarbeit fließen: zuerst Wiedergabe und Streaming, dann Fotografie, Telefonie, generative KI, Spiele und zuletzt die Spezialverbindungen (Satellit, Drohnen, Roboter). Video macht nach der Zählung von Ericsson Ende 2025 etwa 75 % des gesamten mobilen Datenverkehrs aus. 2025 lieferte die Branche etwa 1,26 Milliarden Smartphones (IDC) und rund 196–209 Millionen Fernseher aus (Schätzungen von TrendForce und Omdia). Jedes dieser Geräte decodiert Video in festverdrahteter Hardware.

Wie ich gearbeitet habe

Dies ist eine Sekundärrecherche; ich bin dabei so vorgegangen, wie ich einen Chip oder einen Produktplan prüfen würde. Ich habe Standarddokumente und Sitzungsberichte (JVET, MPEG, JPEG, AOMedia, IETF) gelesen, Engineering-Blogs und Release Notes von Unternehmen, Chip-Paper, Gebührentabellen von Patentpools, Texte zu Exportkontrollen und die Presse. Jede Zahl trägt ihr Datum. Wo ich eine Zahl selbst ableite, etwa den Rechenaufwand pro Pixel aus den Tabellen eines Papers, sage ich das und kennzeichne sie mit **Schätzung**. Angaben, die ich nicht an einer Primärquelle prüfen konnte, sind mit **Anbieterangabe** gekennzeichnet. Beträge sind in US-Dollar angegeben. Stand ist der 12. November 2026.

Ich vergleiche Gleiches mit Gleichem, wo ich kann, und sage es, wo ich es nicht kann. Kompressionsgewinne bedeuten wenig ohne vier Angaben: die Referenz (Gewinn gegenüber was?), die Metrik (PSNR, VMAF oder ein Zuschauerpanel), Inhalt und Auflösung sowie die Kosten der Decodierung. Der größte Teil der Meinungsverschiedenheiten über neuronale Codecs verschwindet, sobald diese vier genannt sind.

KAPITEL 02

WIEDERGABE UND STREAMING

02 Wiedergabe und Streaming: neuronal an beiden Enden, Standard in der Mitte

Was Zuschauer empfangen, ist weiterhin ein Standard-Bitstrom

Die größten Videodienste haben einen Codec-Wechsel abgeschlossen und beginnen den nächsten, und bei keinem von beiden spielt ein neuronaler Decoder eine Rolle. Im Dezember 2025 meldete Netflix, dass etwa 30 % seines Streamings auf AV1 entfielen, den nach AVC am zweithäufigsten genutzten Codec des Dienstes. AV1 erreichte die Fernseher erst mit den Hardware-Decodern: Netflix nahm AV1 2019 in seine Gerätezertifizierung auf, streamte ihn ab Ende 2021 auf Fernseher und gibt an, dass 88 % der zwischen 2021 und 2025 zur Zertifizierung eingereichten Geräte mit großem Bildschirm ihn unterstützten, fast alle seit 2023. YouTube gab an, dass mit Stand 19. März 2026 mehr als 90 % seines Katalogs, gewichtet nach Wiedergabezeit, in AV1 verfügbar waren.

AV2, der Nachfolger von AV1 aus der Alliance for Open Media, erreichte am 28. Mai 2026 Version 1.0.0, mit etwa 30 % besserer Kompression als AV1 nach den objektiven Messungen von AOMedia. YouTube plant AV2-Streams im Web „schon ab Anfang 2027“. Das Silizium folgte sofort: Chips&Media stellte im Mai einen AV2-Hardware-Decoder fertig und lizenzierte ihn im Juni an einen ungenannten nordamerikanischen Big-Tech-Konzern; Allegro DVT hatte seine AV2-Decoder-IP bis zum 16. Juni an drei Kunden ausgeliefert; VeriSilicon hatte bis zum 10. Juni eine AV2-fähige Video-IP an mehrere Kunden freigegeben. AV2 ist ein klassischer Codec: Seine Gewinne stammen aus besserer Partitionierung, Prädiktion und Transformation sowie besseren In-Loop-Filtern (einer davon ein „datengetriebener“ Guided Detail Filter), nicht aus einem End-to-End-Netz.

VVC, der aktuelle Standard von ITU und ISO, fand seine erste Heimat im Rundfunk in Brasilien, dessen System DTV+ (TV 3.0) im Juni 2026 rechtzeitig zur Weltmeisterschaft in den kommerziellen Betrieb ging. Im Juli 2026 hatte noch kein in Brasilien verkaufter Fernseher einen eingebauten DTV+-Empfänger; die ersten werden zwischen Ende 2026 und Anfang 2027 erwartet.

Am Encoder: der neuronale Downscaler von Netflix

Das sauberste Beispiel für neuronale Codierung in der Wiedergabe läuft seit 2022 und ändert nichts an dem, was das Gerät eines Zuschauers tun muss. Netflix codiert jeden Titel in mehreren Auflösungen, und der Downscaling-Schritt, der aus einem 4K-Master 1080p, 720p und 540p erzeugt, nutzte traditionell einen festen Filter wie Lanczos. Netflix ersetzte ihn durch ein kleines neuronales Netz, einen Vorverarbeitungsblock gefolgt von einem Skalierungsblock, trainiert darauf, die Qualität zu verbessern, die nach einem Standard-Upscaler auf dem Gerät zu sehen ist. Netflix beschreibt es als Drop-in-Lösung: keine weitere Änderung auf der Encoding- oder der Client-Seite, sodass „Millionen von Geräten, die Netflix-Streaming unterstützen, automatisch profitieren“. In den Präferenztests von Netflix wählten etwa 77 % der Zuschauer das Ergebnis des neuronalen Downscalers Anbieterangabe. Damit die Kosten in der Größenordnung von Netflix tragbar bleiben, „genügten schon wenige Schichten eines neuronalen Netzes“.

Dieses Muster sollte man sich merken. Das Netz läuft einmal pro Titel im Rechenzentrum, wo Rechenleistung billig und planbar ist, und sein Ergebnis ist ein gewöhnlicher AV1- oder HEVC-Stream.

Am Bildschirm: Super-Resolution in der App und im Fernseher

Am anderen Ende reparieren Netze, was Codec und Quelle verloren haben. Am 29. Oktober 2025 kündigte YouTube Super Resolution für seine TV-App an: Unter 1080p hochgeladene Videos, von 240p bis 720p, werden per maschinellem Lernen auf HD hochskaliert, erklärtes Ziel ist 4K. Die hochskalierten Versionen sind gekennzeichnet, Zuschauer können zum Original zurückschalten, und Creator können sich dagegen entscheiden. Meta beschrieb im Oktober 2025, wie es Video-Super-Resolution sowohl auf seinen Servern beim Upload betreibt, auf Standard-x86-CPUs in Partnerschaft mit Intel, als auch auf dem Client bei der Wiedergabe, und stellte fest, dass sich beides ergänzt. TV-Hersteller tun dasselbe in Silizium: Samsungs 8K-Flaggschiff-Prozessor NQ8 AI Gen3 wirbt mit 768 neuronalen Netzen für das Upscaling.

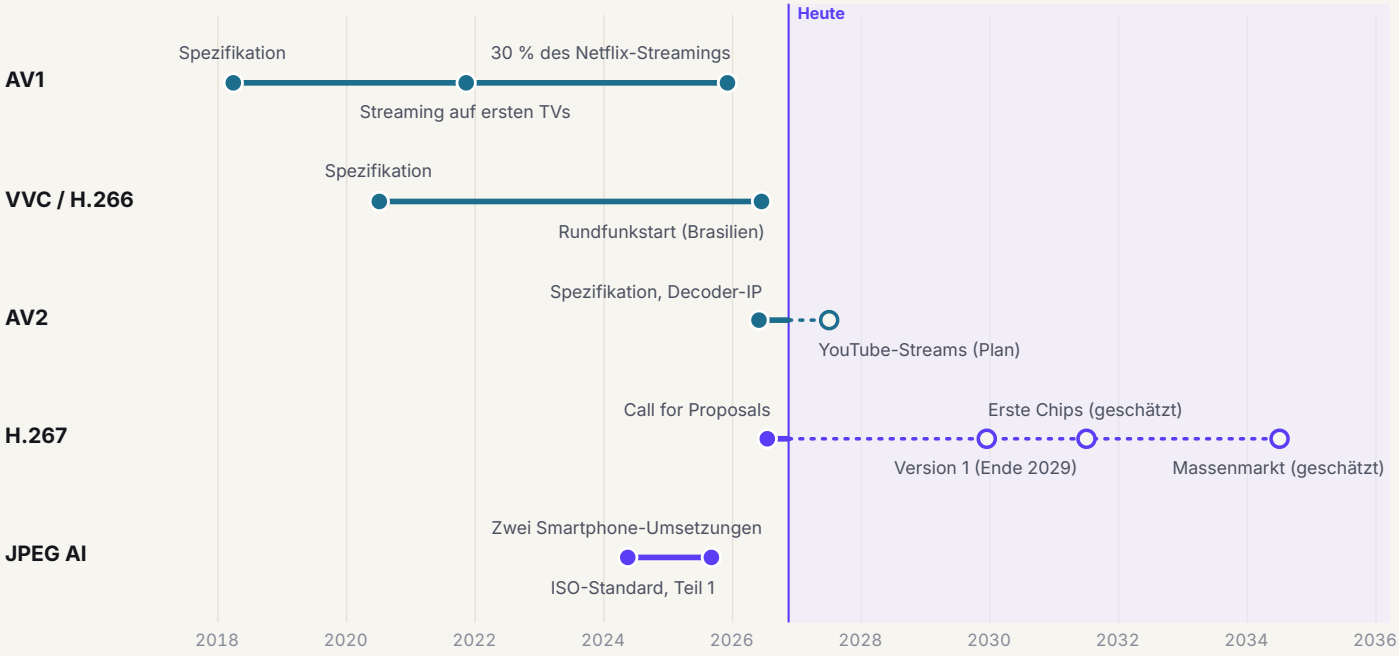
Die Wirtschaftlichkeit hängt davon ab, wo das Netz läuft. Serverseitige Verarbeitung kostet die Plattform Rechenleistung für jedes Video, das sie bearbeitet. Geräteseitige Verarbeitung kostet die Plattform nichts; sie nutzt eine NPU, die der Käufer bereits bezahlt hat. Deshalb kommen die größten Gewinne 2026 aus beidem.

Warum kein Fernseher einen neuronalen Bitstrom decodiert

Ein Wiedergabeformat muss auf jedem Gerät decodierbar sein, das es je abspielen wird, und diese Geräte halten jahrelang. Das macht den Weg vom fertigen Standard bis zum Massenmarkt selbst für klassische Codecs langsam. Die Spezifikation von AV1 war im März 2018 fertig; der Software-Decoder dav1d folgte binnen Monaten; ab Ende 2021 kam der Codec per Streaming auf Fernseher; im Dezember 2025 erreichte er etwa 30 % des Netflix-Streamings, fast acht Jahre nach der Spezifikation. VVC war im Juli 2020 fertig und erreichte im Juni 2026 seinen ersten Rundfunkdienst.

ABB. 03 **Vom fertigen Standard bis zum Massenmarkt vergeht der Großteil eines Jahrzehnts**

Für die Wiedergabe braucht das Gerät einen Hardware-Decoder. AV1 brauchte von der fertigen Spezifikation bis zu 30 % des Netflix-Streamings fast acht Jahre; der erste Standard, der wahrscheinlich neuronale Werkzeuge enthält, ist für Ende 2029 geplant.



Quelle: Netflix Tech Blog (Dez. 2025); AOMedia; Allegro DVT; TV Technology; Fraunhofer IIS zu DTV+ in Brasilien; MPEG-155; JPEG-Komitee; ISO. Hohle Markierungen sind Zielvorgaben oder Schätzungen des Autors.

Ein neuronaler Bitstrom bringt zwei Probleme mit, die klassische Codecs längst gelöst haben: Energie pro Pixel und bitgenaue Decodierung. Beide behandelt Kapitel 8. Zusammen bedeuten sie, dass kein TV- oder Smartphone-Hersteller einen neuronalen Videodecoder in jedes Gerät einbauen wird, bevor ein Standard genau vorgibt, was zu bauen ist.

Was das ändern würde: H.267

Der Standard, der neuronale Werkzeuge auf jeden Bildschirm bringen könnte, ist der Nachfolger von VVC, den ITU-T und ISO/IEC gemeinsam im Joint Video Experts Team (JVET) entwickeln. Sein Name steht noch nicht fest; ich nenne ihn H.267, wie es ein großer Teil der Branche tut.

Im Oktober 2025 wertete JVET die Antworten auf einen Call for Evidence aus: fünf Einreichungen von 16 Unternehmen und Forschungslaboren, mit 20 getesteten Codec-Konfigurationen. Das Gremium fand Belege für bessere Kompression als VVC und bewertete sowohl verbesserte konventionelle Technik als auch Technik auf Basis neuronaler Netze als „interessant“. Neuronale Ansätze zeigten deutliche wahrnehmbare Gewinne, besonders bei 8K-HDR-Inhalten, und kleinere bei Gaming-Inhalten. InterDigital, einer der Einreicher, legte ein Design auf Basis des konventionellen Enhanced Compression Model (ECM) vor und eines, das KI-Werkzeuge über VVC legte.

Der formelle Call for Proposals wurde auf dem Treffen im Juli 2026 in Genf veröffentlicht. Die Haupteinreichungen waren Ende Oktober 2026 fällig, ein Ergänzungspaket ist Mitte Dezember fällig, und die Einreichungen werden auf dem JVET-Treffen in Brisbane im Januar 2027 ausgewertet. JVET erwartet ein erstes Testmodell in der zweiten Jahreshälfte 2027 und eine erste Version des Standards bis Ende 2029 und nennt das Vorhaben „vermutlich einen größeren Aufwand als je zuvor beim Start eines neuen Standards“.

Der klassische Zweig legt die Latte hoch. ECM, gebaut mit konventioneller Signalverarbeitung, spart in der Random-Access-Konfiguration rund 28 % gegenüber VVC. Die Explorationsarbeit von JVET zu neuronalen Netzen meldet In-Loop-Filter, die etwa 6 % bei 10 Tausend Multiply-Accumulate-Operationen pro Pixel sparen, 8 % bei 22 Tausend und 14 % bei einem Betriebspunkt hoher Komplexität. Meine Lesart ist, dass H.267 höchstwahrscheinlich ein klassischer Kern mit einigen neuronalen Werkzeugen dort wird, wo die eingesparten Bits den Decodieraufwand rechtfertigen, und dass die Auswertung im Januar 2027 zeigen wird, wie viele es sind Analyse.

Eine Brücke gibt es bereits. VVC- und HEVC-Streams können Supplemental Enhancement Information (SEI) transportieren, die einen neuronalen Postfilter beschreibt, den die Decoderseite anwenden kann, und die im Oktober 2025 fertiggestellten neuen Ausgaben von VVC und seines SEI-Standards (VSEI) haben diese Nachrichten erweitert. Ein Stream kann angeben, welches Netz nach der Decodierung laufen soll, ohne jeden Decoder zu zwingen, es auszuführen.

Was das für Unternehmen in der Wiedergabe bedeutet

Verkaufen Sie jetzt, was rund um den Codec funktioniert: encoderseitige Vorverarbeitung für Dienste sowie Nachverarbeitungsblöcke und Tuning für TV- und Set-Top-Chips. Bauen Sie vor den 2030er-Jahren kein Wiedergabegeschäft auf einem neuronalen Bitstrom auf, der auf Geräten Dritter decodiert werden muss. Wenn Sie Silizium herstellen, liegt der kurzfristige Umsatz bei AV2- und VVC-Decodern; der neuronale Teil ist ein NPU-Pfad für Bildverbesserung daneben (Kapitel 8).

KAPITEL 03

FOTOGRAFIE UND BILDER

03 Fotografie und Bilder: neuronal in der Kamera, klassisch im Web

Die Kamera-Pipeline ist bereits neuronal

Jedes Foto eines aktuellen Flaggschiff-Smartphones hat neuronale Netze durchlaufen, bevor es komprimiert wird. Qualcomms Snapdragon 8 Gen 2, verbaut in Smartphones ab 2023, führte einen „Cognitive ISP“ ein, der bei Foto- und Videoaufnahmen eine semantische Segmentierung in Echtzeit durchführt und Gesichter, Haare, Himmel und Hintergründe getrennt anpasst; Samsungs Galaxy S23 Ultra war das erste Smartphone, das ihn bei der Selfie-Kamera nutzte. Im Januar 2026 zeigten Chips&Media und Visionary.ai auf der CES, was sie den ersten vollständig KI-basierten ISP nennen: eine durchgängige Bildverarbeitungs-Pipeline, die komplett auf einer NPU läuft, der WAVE-N von Chips&Media, in 16-Bit-Gleitkomma, ohne festverdrahteten ISP, und sich drahtlos aktualisieren lässt. „Bestehende Pipelines schrauben neuronale Blöcke an einen festverdrahteten ISP. Wir ersetzen den konventionellen ISP vollständig“, sagte der CEO von Visionary.ai. Nach der Messe wurde der ISP Herstellern zur Evaluierung angeboten.

Danach wird das Ergebnis als JPEG oder HEIF gespeichert, klassische Formate, die jedes Gerät öffnen kann. Das Muster ist dasselbe wie bei der Wiedergabe: neuronal innerhalb der Pipeline eines Unternehmens, Standard an der Grenze.

JPEG AI: ein fertiger neuronaler Standard, noch ohne Produkt

JPEG AI ist der erste Standard für gelernte Bildcodierung von ISO/IEC und ITU-T (ISO/IEC 6048, ITU-T T.840). Teil 1, das Kerncodiersystem, wurde 2025 veröffentlicht. Das JPEG-Komitee meldet Kompressionsgewinne von 12,5 % bis 27,9 % gegenüber der Intra-Codierung von VVC, über eine Spanne der Decoderkomplexität von 7 bis 216 Tausend Multiply-Accumulate-Operationen pro Pixel. Sein Design umfasst zwei Encoder und drei Decoder, gemeinsam trainiert, sodass ein Empfänger nur einen Decoder implementieren und trotzdem jeden Stream lesen kann. Bis Mai 2024 meldete das Komitee zwei Implementierungen auf Smartphones. Parallel dazu veröffentlichte die IEEE im Dezember 2024 IEEE 1857.11-2024, einen Standard für Bildcodierung auf Basis neuronaler Netze, ausgelegt für NPUs.

Mit Stand November 2026 habe ich keine Smartphone-Kamera, keine Foto-App und keinen Browser gefunden, die JPEG AI standardmäßig schreiben oder lesen.

Was das Web stattdessen übernommen hat: JPEG XL

Das Bildformat, das 2026 tatsächlich an Boden gewonnen hat, ist klassisch. Am 6. Oktober 2026 aktivierte Google die Decodierung von JPEG XL in Chrome 155 standardmäßig, mit einem neuen speichersicheren Decoder in Rust, vier Jahre nachdem Google die experimentelle Unterstützung entfernt hatte. Safari decodiert JPEG XL seit 2023, und Mozilla bewegt sich in dieselbe Richtung. Google begründet das mit besserer Kompression als JPEG, verlustfreier Neukomprimierung bestehender JPEG-Dateien, HDR und progressiver Decodierung.

Die Lehre lässt sich verallgemeinern. Ein Format, das sich überall öffnen lassen muss, gewinnt mit einem schnellen, deterministischen Decoder, der auf jeder CPU läuft und sich speichersicher machen lässt. Ein neuronaler

Bilddecoder braucht eine NPU oder GPU und exakte Arithmetik über alle Geräte hinweg. Neuronale Codierung gewinnt in der eigenen Pipeline eines Geräts und in Systemen, in denen dasselbe Unternehmen codiert und decodiert, etwa bei einem Cloud-Fotodienst, der die komprimierte Datei nie an jemand anderen weitergibt.

Was das für Unternehmen in der Bildverarbeitung bedeutet

Das Geld liegt in der Aufnahme-Pipeline: KI-ISP-IP, für die Bildverarbeitung dimensionierte NPUs und die Tuning-Arbeit, die Kamerahersteller einkaufen. Für Speicherung und Auslieferung ist ein gelernter Bildcodec heute ein Werkzeug für geschlossene Systeme, kein Format, auf das man setzen sollte.

KAPITEL 04



TELEFONIE UND KONFERENZEN

04 Telefonie und Konferenzen: wo neuronale End-to-End-Codecs tatsächlich im Einsatz sind

Audio kam zuerst

Microsoft stellte im Februar 2021 Satin vor, einen KI-gestützten Sprachcodec, und gab an, dass er bereits für alle Zweiergespräche in Teams und Skype genutzt werde. Satin liefert Super-Wideband-Sprache mit bis zu 16 kHz Audiobandbreite bei 6 kbps, wo der frühere Skype-Codec Silk bei dieser Rate schmalbandig war. Die eingesparte Bitrate fließt in Redundanz, sodass Anrufe Paketverluste besser überstehen.

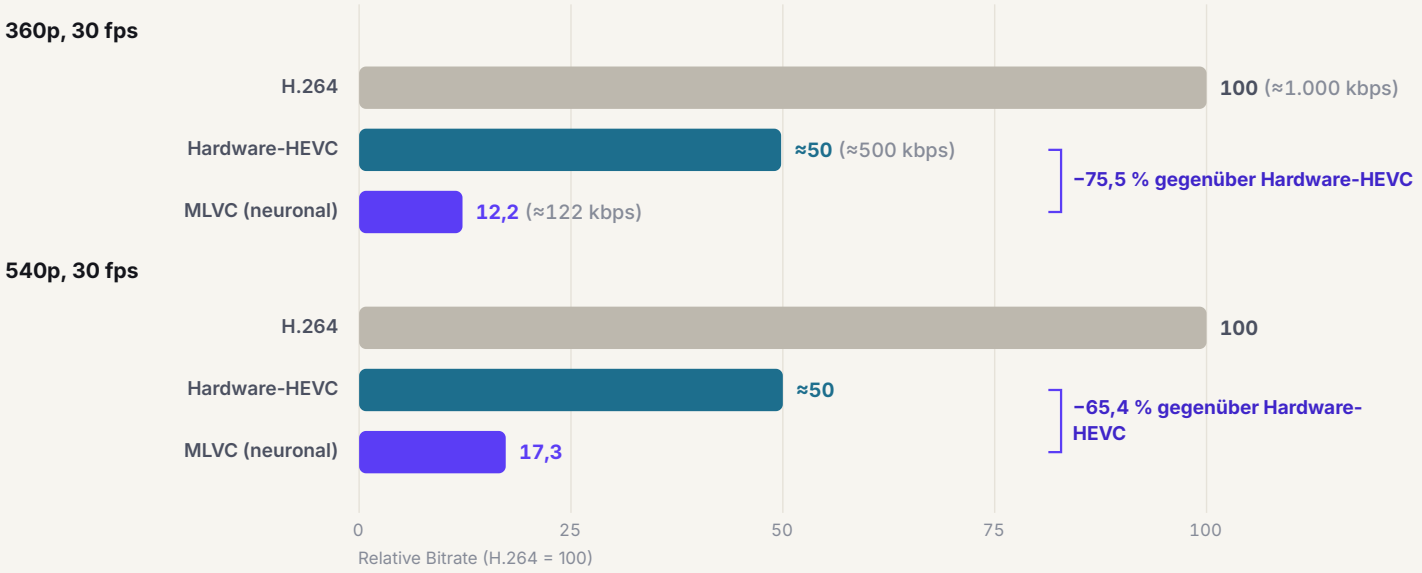
MLVC: ein neuronaler Videocodec in echten Anrufen

Am 24. Juli 2026 veröffentlichte Microsoft MLVC, seinen Videocodec auf Basis maschinellen Lernens, als Open Source unter der MIT-Lizenz, mit Modellcode, trainierten Gewichten, Trainingsskripten und den Konvertierungswerkzeugen für verschiedene NPUs. MLVC ist die Produktversion der DCVC-Familie, die Microsoft Research seit 2021 veröffentlicht. Microsoft zufolge wird MLVC in Teams für Peer-to-Peer-Anrufe eingeführt, mit Live-Telemetrie, A/B-Tests und den konventionellen Codecs als Rückfallebene.

MLVC läuft mit 540p und 360p bei 30 Bildern pro Sekunde und ist so ausgelegt, dass er auf Apples Neural Engine sowie auf NPUs von Qualcomm und Intel nicht mehr als die Hälfte der NPU belegt; die Arbeit teilt er zwischen CPU und NPU auf. In subjektiven Tests nach ITU-T P.910 auf Microsofts eigenem Video Conferencing Dataset brauchte er bei 360p 87,8 % weniger Bitrate als H.264 und 75,5 % weniger als Hardware-HEVC, bei 540p 82,7 % und 65,4 % weniger Anbieterangabe. Bei 360p sind das rund 122 kbps, wo H.264 etwa 1 Mbps braucht.

ABB. 04 **Bei Anrufen schlägt ein neuronaler Codec Hardware-HEVC um 65–76 %**

Benötigte Bitrate für gleiche bewertete Qualität auf Microsofts Testdatensatz für Videokonferenzen, H.264 = 100. MLVC läuft auf NPUs von Laptops und Smartphones mit bis zu 540p30 und belegt höchstens die Hälfte der NPU.



Quelle: Microsoft, Veröffentlichung von MLVC (24. Juli 2026): subjektive Tests nach ITU-T P.910 auf dem eigenen Video Conferencing Dataset. HEVC-Balken aus den angegebenen Einsparungen abgeleitet. Anbieterangaben, keine unabhängigen Messungen.

Zwei Dinge relativieren die Schlagzeile. Die 88 % beziehen sich auf H.264; gegenüber dem Hardware-HEVC, das dieselben Laptops bereits enthalten, sind es 65–76 %, immer noch viel. Und MLVC hat Kompression geopfert, um einsetzbar zu sein: Eine unabhängige Analyse stellt fest, dass eine perzeptuell trainierte Variante von DCVC-RT beim selben Maß 81,9 % erreicht, wo MLVC auf 75,5 % kommt. Der Preis für identische Decodierung auf fremder Hardware lag bei etwa sechs Prozentpunkten.

Warum Anrufe das natürliche Zuhause sind

Ein Anruf ist ein geschlossenes System. Beide Endpunkte nutzen dieselbe Anwendung, handeln zu Beginn des Anrufs aus, was sie können, und können mitten im Gespräch auf einen Standard-Codec ausweichen. Die Auflösungen sind bescheiden, die Verbindungen oft schlecht, und die Laptops an beiden Enden kommen inzwischen mit NPUs, wie Smartphones schon seit Jahren. Microsoft setzt 40 TOPS NPU-Leistung als Messlatte für seine KI-PCs, und Apple gibt die Neural Engine des M4 mit 38 TOPS an. Jede Bedingung, die einen neuronalen Codec bei der Wiedergabe blockiert, fehlt hier.

Die Gegenbeispiele: Opus und MLow

Nicht jede Telefonie-Plattform hat sich so entschieden, und die Gründe sind lehrreich.

Opus, der offene Audiocodec in WebRTC und den meisten Telefonie-Apps, nahm in Version 1.5 (März 2024) Deep Learning auf, zum ersten Mal mit Netzen, die das Signal selbst verarbeiten oder erzeugen. Allerdings nur als Optionen: decoderseitige Verbesserung (LACE und NoLACE), die keine Änderung am Bitstrom erfordert, eine neuronale Verschleierung von Paketverlusten und ein Deep-Redundancy-Verfahren (DRED); alle sind standardmäßig aus, damit alte Geräte nicht belastet werden. Opus 1.6 brachte eine neuronale Bandbreitenerweiterung und ein dreimal kleineres DRED-Modell und bleibt voll kompatibel mit dem Standard von 2012. Der Ansatz lautet „im Standard, optional und abwärtskompatibel".

Meta ging noch weiter zurück. Im Juni 2024 führte das Unternehmen MLow ein, einen Sprachcodec für niedrige Bitraten auf Basis klassischer Signalverarbeitung statt eines Netzes, heute in allen Anrufen in Instagram und Messenger und in der Einführung bei WhatsApp. Metas Gründe waren seine Geräte: Mehr als 20 % seiner Anrufe werden auf ARMv7-Smartphones geführt, und täglich laufen Dutzende Millionen WhatsApp-Anrufe auf Smartphones, die mehr als zehn Jahre alt sind. Neuronale Codecs, auch Metas eigener EnCodec, laufen nur auf teuren Smartphones zuverlässig, schrieb Meta. MLow liefert bei 6 kbps etwa die doppelte Qualität von Opus und braucht dabei 10 % weniger Rechenleistung.

Für einen Dienst, dessen Nutzer von Flaggschiff-Laptops bis zu zehn Jahre alten Smartphones reichen, entscheidet das Komplexitätsbudget. Neuronale Codecs kommen zuerst dort zum Einsatz, wo der Gerätebestand neu ist und der Anbieter beide Enden kontrolliert.

Was das für Unternehmen in der Kommunikation bedeutet

Unternehmenskonferenzen, Telepräsenz, Remote-Assistenz, Zuspielstrecken und private Netzwerke sind die Orte, an denen ein neuronaler End-to-End-Codec heute ein Produkt ist, vorausgesetzt, beide Enden gehören Ihnen. Planen Sie von Anfang an drei Dinge ein: einen Ausweich-Codec, Decodiertests auf jeder NPU, die Sie unterstützen, und eine Qualitätsbewertung mit Ihren eigenen Inhalten statt mit dem Datensatz eines Anbieters.

KAPITEL 05

GENERATIVE KI

05 Generative KI: der größte Einsatz, den niemand Codec nennt

Jedes Voice- und Videomodell läuft auf einem gelernten Codec

Der größte Einsatz neuronaler End-to-End-Kompression steckt in keinem Codec-Standard. Er steckt in generativer KI. Ein Voice-Modell hört keine Wellenformen und erzeugt sie auch nicht direkt: Es liest und schreibt die diskreten Tokens eines gelernten Audiocodecs. Mimi von Kyutai, 2024 veröffentlicht, wandelt Sprache mit 12,5 Frames pro Sekunde und 1,1 kbps in Tokens um, weniger als jeder Telefonie-Codec, und vereint semantische und akustische Information in einem Datenstrom. Metas EnCodec, 2022 veröffentlicht, ist ein weiterer; Metas eigene Ingenieure beschrieben solche Codecs später als hervorragend bei niedrigen Bitraten, aber zu schwer für ältere Smartphones (Kapitel 4).

Videogenerierung funktioniert genauso. Ein Diffusions- oder Transformer-Modell generiert im Latentraum eines gelernten Video-Autoencoders (VAE), und der Decoder des VAE verwandelt das Ergebnis in Pixel. Der VAE von LTX-Video komprimiert Video im Verhältnis 1:192: Jedes latente Token steht für einen Block von 32×32 Pixeln über 8 Frames, mit 128 Kanälen. Sein Decoder übernimmt auch den letzten Entrauschungsschritt und erzeugt Details damit teilweise, statt sie nur zu rekonstruieren.

Warum das keine Distributionsformate sind

Diese Codecs laufen bei jeder Anfrage an ein Voice- oder Videomodell, aber keiner von ihnen ist ein Format, das jemand anderes decodiert. Jedes Modell hat seinen eigenen Latentraum; es gibt keine Ratenregelung, keinen Konformitätstest und keine Garantie, dass ein anderer Decoder dasselbe Bild erzeugen würde. Sie sind auf Generierungsqualität ausgelegt, nicht auf Wiedergabetreue bei gegebener Bitrate, und sie decodieren auf GPUs im Re-

chenzentrum. Ihr Erfolg zeigt, dass gelernte Kompression in großem Maßstab funktioniert, wenn einem Unternehmen Encoder, Decoder und Hardware gehören.

Was das für Unternehmen bedeutet

Zwei praktische Punkte. Erstens ist die Effizienz des Tokenizers ein direkter Kostenposten für jeden, der generatives Video oder generierte Sprache anbietet: Bei einem Transformer, der Attention über alle Tokens berechnet, wächst der Rechenaufwand schneller als die Zahl der Tokens, sodass eine stärker komprimierende latente Darstellung die Generierungskosten überproportional senkt Analyse. Codec-Ingenieure, die Latent-Design, Entropiemodelle und Decodereffizienz verstehen, sind für Teams im Bereich generativer KI wertvoll. Zweitens verschiebt sich die Grenze: Standardisierungsgremien haben mit der Arbeit an Codierung für Maschinen statt für Menschen begonnen. MPEGs Video Coding for Machines (ISO/IEC 23888-2) erreichte Anfang 2026 das Stadium des internationalen Standardentwurfs und soll Ende 2026 final sein, und JPEG AI wurde von Anfang an darauf ausgelegt, Aufgaben der Computer Vision ebenso zu dienen wie dem menschlichen Betrachten.

KAPITEL 06

SPIELE UND GRAFIK

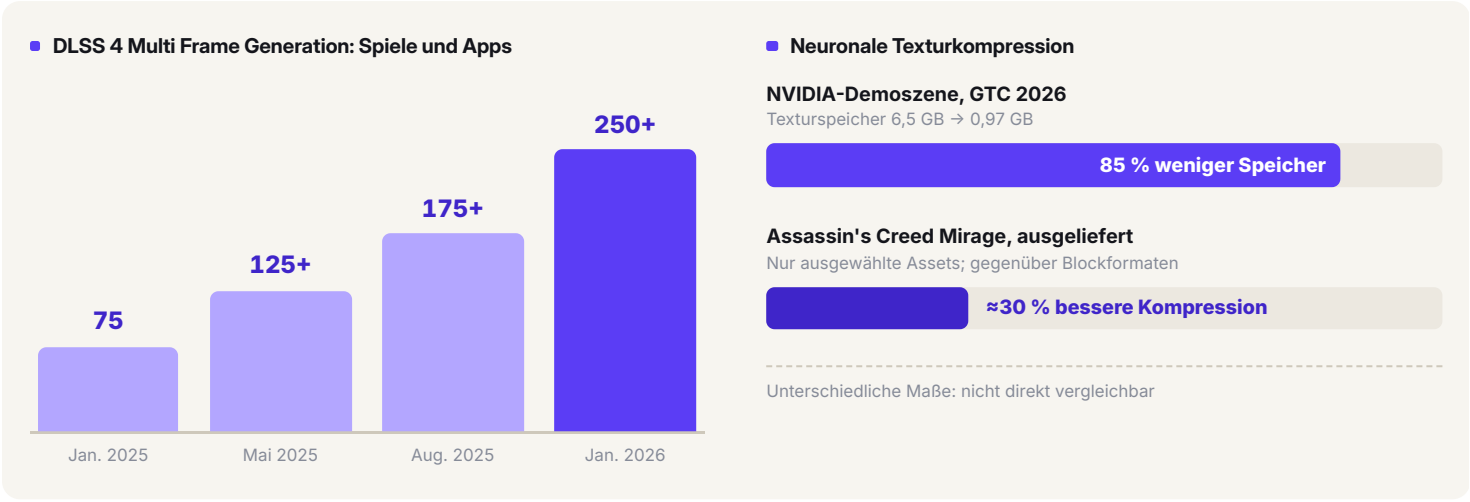
06 Spiele und Grafik: neuronale Rekonstruktion ist etabliert, neuronale Assets stehen am Anfang

Weniger Pixel rendern und ein Netz den Rest ergänzen lassen

In der Echtzeitgrafik rekonstruiert bereits ein Netz den größten Teil dessen, was Spieler sehen. NVIDIAs DLSS rendert weniger Pixel und weniger Frames und erzeugt den Rest mit Netzen. DLSS 4 mit Multi Frame Generation startete im Januar 2025 mit Unterstützung in 75 Spielen und Apps; NVIDIA zählte bis Mai 2025 mehr als 125, bis August mehr als 175 und bis Januar 2026 mehr als 250, als das Unternehmen DLSS 4.5 mit einem Transformer-Modell der zweiten Generation für Super-Resolution ankündigte. Das ist faktisch ein Decoder, nur heißt er nicht so: Der „Bitstrom“ ist ein Frame in niedriger Auflösung plus Bewegungsdaten, und das Netz rekonstruiert das volle Bild.

ABB. 05 **Mit dem Stack aus einer Hand verbreitet sich neuronale Rekonstruktion schnell**

Links: Spiele und Apps mit Unterstützung für NVIDIAs DLSS 4 Multi Frame Generation. Rechts: neuronale Texturkompression in einer Demoszene und in einem ausgelieferten Spiel, das sie wegen der Laufzeitkosten nur für ausgewählte Assets nutzte.



Quelle: NVIDIA (Januar 2025 bis Januar 2026; Demo auf der GTC 2026); Ubisoft La Forge (Februar 2026). Anbieterangaben; die beiden Texturwerte messen Unterschiedliches.

Neuronale Texturen sind ausgeliefert, wenn auch selektiv

Texturen werden in festen Blockformaten (BCn) komprimiert, die GPUs in Hardware decodieren. Neuronale Texturkompression ersetzt einen Satz von Materialtexturen durch eine kompakte latente Repräsentation und ein kleines Netz, das Texel bei Bedarf rekonstruiert. Im Februar 2026 beschrieb Ubisofts Labor La Forge, wie sie mit Assassin's Creed Mirage ausgeliefert wurde, einem Spiel von 2023, mit etwa 30 % höherer effektiver Kompression als die konventionellen Formate bei gleicher visueller Qualität Anbieterangabe. Sie kam nicht überall zum Einsatz: „Die Nutzung war durch die Laufzeitleistung begrenzt“, deshalb wandte Ubisoft sie auf eine Teilmenge der Assets mit hohen Instanzzahlen und großem Speicherdruck an.

NVIDIAs RTX Neural Texture Compression SDK, als Beta auf GitHub verfügbar, zeigt die Obergrenze. In einer Demonstration auf der GTC 2026 brauchte eine Villa-Szene, die mit BCn 6,5 GB Texturspeicher belegte, mit neuronalen Texturen 970 MB, 85 % weniger Anbieterangabe. Das SDK bietet zwei Wege: beim Laden nach BCn decodieren, was Speicherplatz und Downloadgröße spart, aber keinen Grafikspeicher, oder beim Texturzugriff decodieren, was den Speicher spart, aber in jedem Frame Shader-Zeit kostet.

Die API kommt

Grafik-APIs erhalten die Matrixoperationen, die neuronale Decodierung in Shadern braucht. Microsoft hat Shader Model 6.9 im Februar 2026 ausgeliefert; die Cooperative-Vector-Funktion, die 2025 als Vorschau erschien, ging in einem einheitlichen Linear-Algebra-Design für ein späteres Shader Model auf, und DirectX Linear Algebra ging im April 2026 in die öffentliche Vorschau. Wenn das in jedem Treiber steckt, kann ein neuronaler Texturdecoder aus einem einzigen Codepfad auf der GPU jedes Herstellers laufen.

Was das für Unternehmen bedeutet

Spiele liefern neuronale Decodierung aus, weil die Bedingungen eines geschlossenen Systems erfüllt sind: eine Engine, eine Asset-Pipeline, Tensor-Hardware in der GPU und ein Ausweichformat. Die offene Arbeit liegt bei Werkzeugen, Asset-Pipelines, die neuronale Materialien codieren, herstellerübergreifenden Laufzeitumgebungen und vor allem bei Mobilgeräten, wo die Energiekosten pro decodiertem Texel entscheiden.

KAPITEL 07

07

SPEZIALVERBINDUNGEN

Spezialverbindungen: Satelliten, Drohnen und Roboter

Hier passt neuronale End-to-End-Codierung auf dem Papier am besten. Ein Betreiber kontrolliert beide Enden der Verbindung, die Bandbreite ist knapp und verlustbehaftet, die Inhalte sind eng genug, um dafür zu trainieren, und ein wahrnehmungstreues oder sogar nur semantisch treues Bild genügt oft. Verlusttolerante neuronale Codecs für solche Verbindungen sind ein aktives Forschungsgebiet, und serverseitige Restaurierung von Feeds mit niedriger Bitrate wird genutzt.

Verglichen mit Video für Endkunden sind sie zugleich ein kleiner und zunehmend regulierter Markt Analyse. In China überarbeitet die Bekanntmachung Nr. 49 von 2026 des Handelsministeriums und zweier weiterer Behörden, veröffentlicht am 19. Oktober und gültig ab 18. November 2026, die Exportkontrollen für 13 Kategorien drohnenbezogener Güter. Unter den Technologieeinträgen erfasst einer unbemannte Bodenplattformen, die laut Fachpresse teils über Schwellenwerte wie 75 % semantische Bildkompression und Bildrekonstruktion innerhalb von 150 Millisekunden definiert sind. Semantische Kompression ist damit selbst zu einem Parameter der Exportkontrolle geworden. In den USA hat die Federal Communications Commission am 22. Dezember 2025 im Ausland hergestellte Drohnen und ihre kritischen Komponenten in ihre Covered List aufgenommen, sodass neue ausländische Modelle keine Gerätezulassung mehr erhalten können.

Das Fazit: eine solide Nische für ein spezialisiertes Unternehmen mit Industrie- oder Regierungskunden und der Kapazität, Export-Compliance auf beiden Seiten zu bewältigen, aber kein Massenmarkt und nicht der Ort, an dem sich zeigt, ob neuronale Codierung in großem Maßstab funktioniert.

KAPITEL 08

SILIZIUM

08 Der Blick aufs Silizium: Energie, Rechenaufwand, Speicher und Exaktheit

In diesem Kapitel zählt mein eigener Hintergrund am meisten. Ich habe viele Jahre Chipentwicklung und Management im Fabless-Bereich hinter mir, darunter den Full-Custom-Entwurf von ASIC-Datenpfaden und Multimedia-Kamerachip-Projekte wie Beschleuniger für Videocodecs, und die Fragen, die ein Chipteam zu einem neuen Codec stellt, drehen sich nicht um die BD-Rate. Sie lauten: wie viele Joule pro Pixel, wie viele Byte DRAM-Verkehr pro Pixel, wie viele Gatter, und kann jeder Decoder der Welt exakt dieselben Bits berechnen? Bei allen vier steht ein neuronaler Decoder 2026 dort, wo Hardware-HEVC vor seinen ersten Chips stand: vielversprechend in Software, teuer in Silizium.

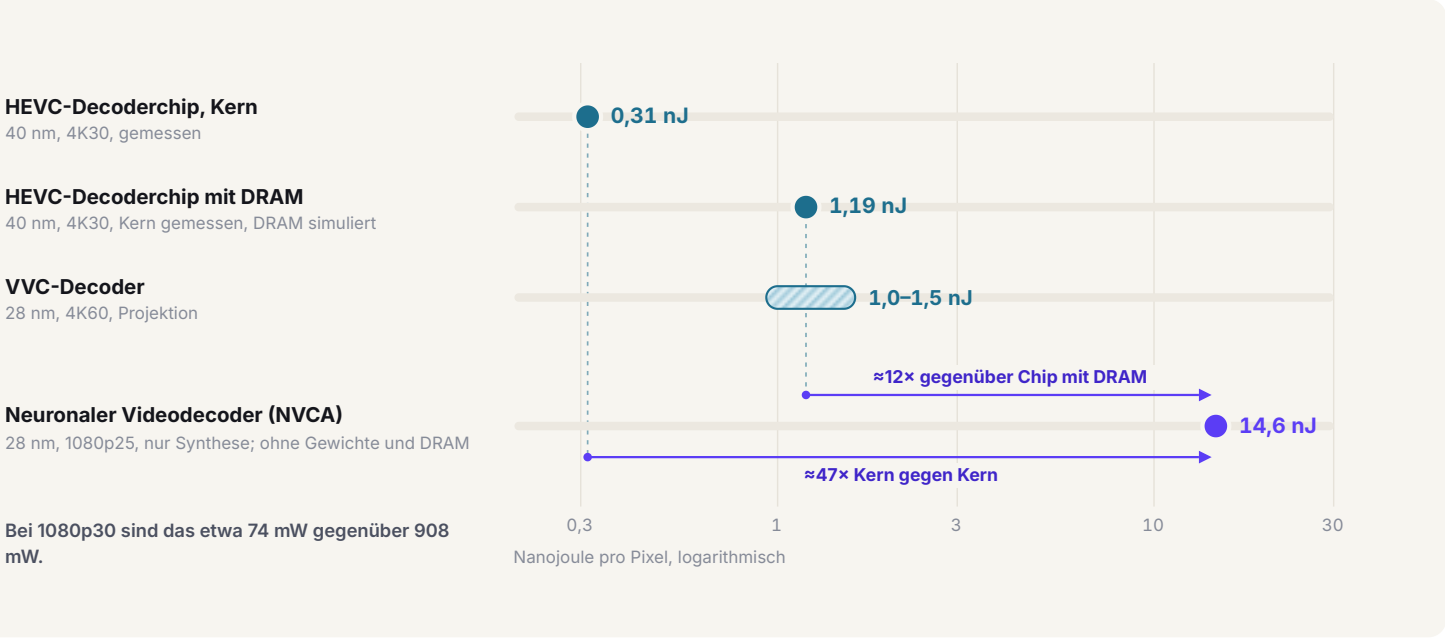
Energie pro Pixel: eine Größenordnung Abstand

Nur wenige neuronale Decoder wurden als Chips entworfen, daher stützt sich der Vergleich auf eine Handvoll veröffentlichter Designs. Der Bezugspunkt auf der klassischen Seite ist der HEVC-Decoderchip des MIT (Tikekar et al., ISSCC 2013 und IEEE JSSC 2014): 40 nm, 1,77 mm², 4K-Decodierung mit 30 Bildern pro Sekunde bei 200 MHz. Sein Kern verbraucht 78 mW oder 0,31 nJ pro Pixel; mit simulierter DRAM-Leistung gerechnet, braucht das System 1,19 nJ pro Pixel. Für VVC sind Produktionswerte proprietär; eine Analyse in Electronic Design von 2026, die den MIT-Chip auf 28 nm skaliert und die höhere Komplexität von VVC ansetzt, projiziert 1–1,5 nJ pro Pixel für einen 4K60-Decoder mit etwa einer Million Gattern.

Auf der neuronalen Seite ist NVCA (Zhang et al., DATE 2024) nach der Zählung von Electronic Design das einzige veröffentlichte ASIC-Design eines neuronalen End-to-End-Videodecoders, synthetisiert im 28-nm-Prozess von TSMC mit 5,01 Millionen Gattern, ohne die Gewichte, die außerhalb des Chips liegen. NVCA decodiert 1080p mit 25 Bildern pro Sekunde bei 760 mW, also 14,6 nJ pro Pixel.

ABB. 06 **Energie pro decodiertem Pixel: Festverdrahtete Decoder liegen eine Größenordnung vorn**

Das einzige veröffentlichte Chipdesign eines neuronalen Videodecoders braucht pro Pixel etwa das 12-Fache der Energie eines gemessenen HEVC-Decoderchips samt DRAM und etwa das 47-Fache von dessen Kern allein. Gewichts- und DRAM-Verkehr des neuronalen Designs sind nicht mitgezählt.



Quelle: Tikekar et al., IEEE JSSC 2014 (MIT); NVCA, Zhang et al., DATE 2024 (arXiv 2312.10716); die VVC-Spanne ist eine Projektion von Electronic Design (2026). Der HEVC-Chip nutzt einen älteren Fertigungsprozess (40 nm gegenüber 28 nm), was das neuronale Design begünstigt.

Das ist etwa das 12-Fache des HEVC-Chips samt DRAM und etwa das 47-Fache seines Kerns allein. Der Vergleich begünstigt NVCA in dreifacher Hinsicht: Der HEVC-Chip nutzt einen älteren Fertigungsprozess, der Wert von NVCA lässt den Gewichts- und DRAM-Verkehr weg, und NVCA ist ein Syntheseergebnis statt gemessenen Siliziums. Bei 1080p30 beträgt der Unterschied etwa 74 mW gegenüber rund 900 mW: ein Decoder, den niemand bemerkt, gegenüber einem, der sich im Akku- und Wärmebudget eines Smartphones bemerkbar macht.

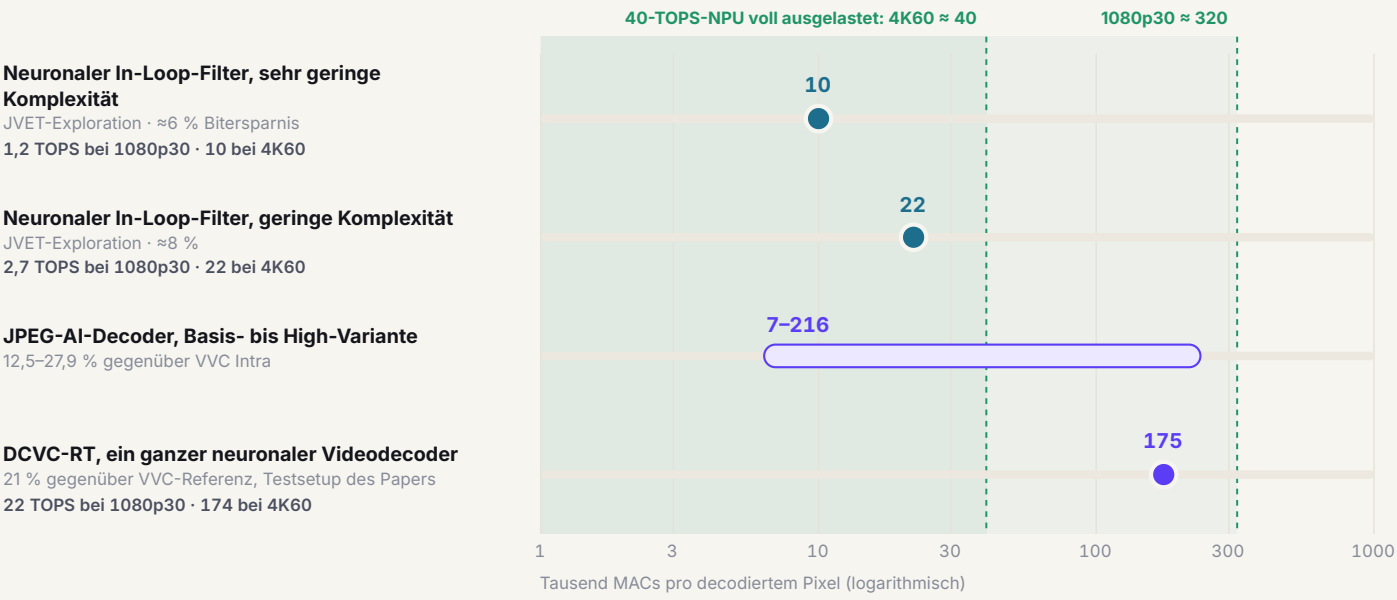
Rechenaufwand pro Pixel: was ein neuronaler Decoder von einer NPU verlangt

Der schnellste veröffentlichte neuronale Videocodec, den ich gefunden habe, ist DCVC-RT von Microsoft Research (Jia et al., CVPR 2025). Er codiert und decodiert 1080p mit 125 bzw. 113 Bildern pro Sekunde auf einer NVIDIA-A100-GPU, spart dabei im Testsetup des Papers 21 % Bitrate gegenüber dem Referenzencoder von VVC und nutzt Ganzzahlarithmetik, damit verschiedene Geräte konsistent decodieren. Sein Zusatzmaterial listet die Komplexität jedes Moduls auf. Addiere ich die vier Module, die ein Decoder ausführen muss (Decoder, Feature-Extraktor, Entropiemodell und Rekonstruktion), komme ich auf etwa 175 Tausend Multiply-Accumulate-Operationen pro decodiertem Pixel **Schätzung**.

Bei 1080p30 sind das etwa 11 Billionen Multiply-Accumulate-Operationen pro Sekunde oder etwa 22 TOPS bei zwei Operationen pro MAC, vor jedem Auslastungsverlust; 4K60 braucht das Achtfache **Schätzung**. Zur Einordnung: Microsoft definiert seine KI-PCs über eine NPU mit mindestens 40 TOPS, und Apple gibt die Neural Engine des M4 mit 38 TOPS an. Das sind Spitzenwerte, und das Gerät teilt sie mit allem anderen, was es tut.

ABB. 07 **Ein neuronales Werkzeug kostet Zehntausende MACs pro Pixel, ein ganzer neuronaler Decoder Hunderttausende**

Multiply-Accumulate-Operationen des Decoders pro Pixel, logarithmisch, verglichen mit dem, was eine voll ausgelastete 40-TOPS-NPU bei 4K60 und bei 1080p30 leisten könnte. Die Gewinne beziehen sich auf unterschiedliche Referenzen und sind nicht direkt vergleichbar.



Quelle: JVET über den Vorsitzenden von SC 29 (April 2026); JPEG-Komitee (2024); DCVC-RT, Zusatzmaterial zu CVPR 2025, Summe der Decodermodule durch den Autor; Microsofts Definition des Copilot+ PC. Rechnung des Autors: zwei Operationen pro MAC, volle Auslastung.

Die Stufenleiter in Abbildung 7 zeigt, wie sich das Feld sinnvoll lesen lässt. Neuronale Werkzeuge, die Zehntausende MACs pro Pixel kosten, etwa die In-Loop-Filter von JVET (etwa 6 % Einsparung bei 10 Tausend, 8 % bei 22 Tausend), sind in einem künftigen Standard-Decoder plausibel. JPEG AI deckt die Spanne ab, von 7 bis 216 Tausend MACs pro Pixel für 12,5–27,9 % gegenüber VVC Intra. Ein ganzer neuronaler Videodecoder mit ein paar hunderttausend MACs pro Pixel gehört auf eine NPU, bei bescheidener Auflösung, in ein geschlossenes System. Genau dort steht MLVC: 540p30, bei höchstens der Hälfte der NPU.

Speicherverkehr, nicht MACs, ist oft die eigentliche Grenze

Die Autoren von DCVC-RT stellten fest, dass nicht die Arithmetik die Geschwindigkeit begrenzte, sondern die „operativen Kosten“: Speicherein- und -ausgabe und die Zahl der Funktionsaufrufe. Vor allem deshalb entfernten sie explizite Bewegungsmodule und die Multiskalenverarbeitung. Die Entwickler von NVCA bauten einen speziellen Datenfluss, um den externen Speicherverkehr durch Bewegungs- und Residual-Features zu verringern. Ein Chipdesigner wird das Muster wiedererkennen. Im HEVC-Chip des MIT entfallen etwa drei Viertel der Systemenergie auf DRAM (0,88 von 1,19 nJ pro Pixel), selbst nach einem Cache, der die DRAM-Bandbreite um 67 % senkte Analyse. Ein neuronaler Decoder, der große mehrkanalige Feature-Maps von Frame zu Frame vorhält, verschärft das Speicherproblem, statt es zu lindern.

Exaktheit: Jeder Decoder muss dieselben Bits berechnen

Entropiecodierung funktioniert nur, wenn das Wahrscheinlichkeitsmodell des Decoders exakt mit dem des Encoders übereinstimmt. In einem klassischen Codec ist dieses Modell eine Tabelle oder ein kleiner Satz von Ganzzahlregeln. In einem neuronalen Codec ist es ein Netz, und eine Gleitkommaabweichung zwischen zwei Geräten,

ausgelöst durch einen anderen Kernel, eine andere Akkumulationsreihenfolge oder eine andere NPU, kann den Decoder die falschen Symbole lesen lassen, sodass er den Stream ab diesem Punkt verliert.

Jeder, der einen neuronalen Codec ausgeliefert oder standardisiert hat, musste dieses Problem technisch ausräumen. DCVC-RT wandelt sein Modell in Ganzzahlen um, damit es geräteübergreifend konsistent codiert. MLVC nutzt ein Hyperprior-Design mit gemeinsamer Skala, um plattformübergreifende Abweichungen zu überstehen, was wahrscheinlich ein Teil des Grundes ist, warum er gegenüber DCVC-RT Kompression einbüßt Analyse. Die Explorationsarbeit von JVET zu neuronalen Netzen lässt jedes Netz über SADL laufen, eine kleine Bibliothek mit 16-Bit-Ganzzahl-Inferenz. JPEG AI definiert Konformitätspunkte im latenten, im Residual- und im Pixelbereich. Für ein Chipteam ist das dieselbe Disziplin wie bei den Ganzzahltransformationen von HEVC: Der Standard muss die Arithmetik exakt festlegen, was das Netzdesign einschränkt und herstellerspezifische Abkürzungen ausschließt.

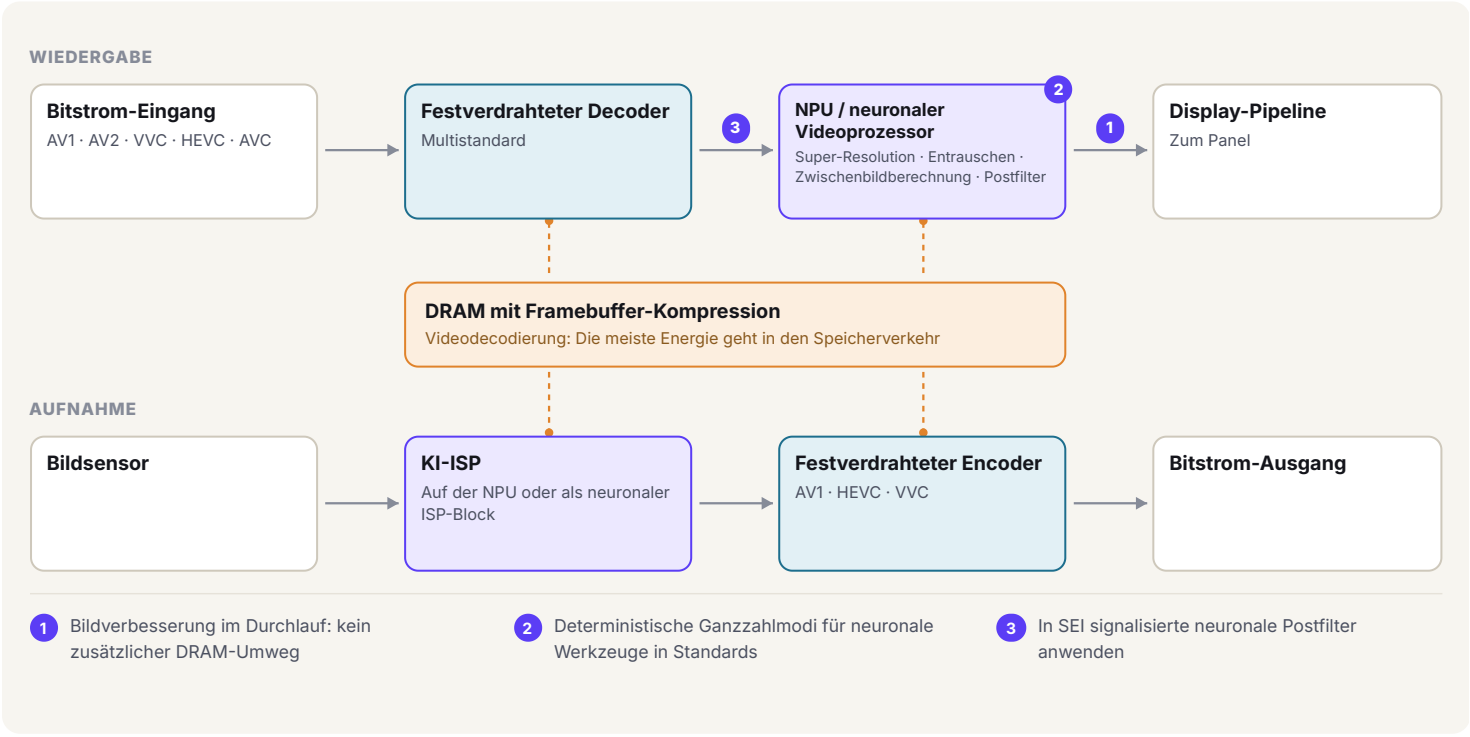
Die Fläche ist nicht das Hindernis, sondern Verifikation und Lebensdauer

Die 5 Millionen Gatter von NVCA gegenüber etwa einer Million für einen projizierten VVC-Decoder wirken viel, aber auf einem modernen Smartphone- oder TV-SoC, das bereits eine NPU mit mehreren TOPS trägt, entscheiden ein paar Millionen Gatter nicht über das Design. Was wirklich Geld kostet, sind die Verifikation gegen eine Konformitäts-Testsuite, die Aufgabe, einen programmierbaren Decoder über Softwareupdates hinweg ein Geräteleben lang korrekt zu halten, und die Leistungsaufnahme bei 4K60 HDR. Diese Kosten ergeben erst Sinn, wenn ein Standard festlegt, was gebaut werden muss.

Die Architektur, die jetzt gewinnt: festverdrahteter Codec plus NPU

ABB. 08 **Die Aufteilung in Serienchips: festverdrahtete Hardware für den Bitstrom, eine neuronale Einheit für die Pixel**

Ein Medien-SoC von 2026. Der Codec-Block decodiert jeden Standard mit einem Bruchteil eines Nanojoule pro Pixel; die NPU übernimmt, was Netze gut können und Standards nicht festlegen. Der DRAM-Verkehr ist der entscheidende Kostenfaktor.



Quelle: Architektursicht des Autors auf Basis von Produktangaben von Tencent (Canghai V2), Chips&Media, Allegro DVT und Samsung (2026).

Das Silizium, das sich 2026 verkauft, kombiniert einen festverdrahteten Multistandard-Codec mit einer neuronalen Einheit daneben. Der Codec-Block verarbeitet AV1, AV2, VVC, HEVC und AVC mit einem Bruchteil eines Nanojoule pro Pixel. Die neuronale Einheit übernimmt, was Netze gut können und Standards nicht festlegen: Entrauschen und den ISP bei der Aufnahme, Super-Resolution, Zwischenbildberechnung und Bildverbesserung bei der Anzeige. Beide teilen sich die Framebuffer-Kompression und die DRAM-Zeitplanung, weil der Speicherverkehr der entscheidende Kostenfaktor ist.

Der Markt zeigt die Aufteilung. Von Tencents Videocodec-ASIC Canghai (沧海), seit 2019 im eigenen Haus entwickelt, sind mehr als 100.000 Chips der ersten Generation in Tencent Cloud und den eigenen Live-, On-Demand- und Cloud-Gaming-Diensten im Einsatz; die zweite Generation, die laut Ankündigung vom Mai 2026 in Betrieb genommen ist und für den Einsatz in der zweiten Jahreshälfte in die Serienfertigung geht, unterstützt H.265 und H.266 in einer Architektur mit mehr als 10 % bzw. 30 % besserer Kompression als der Vorgänger, etwa doppeltem Durchsatz und Bildverbesserung wie Super-Resolution auf dem Chip. Chips&Media verkauft Codec-IP und eine videoorientierte NPU, WAVE-N, nebeneinander. Allegro DVT verkauft AV2-Decoder und hat seine erste IP für neuronale Videoverarbeitung ins Programm genommen. Samsungs TV-Prozessoren kombinieren Standard-Decoder mit Hunderten von Upscaling-Netzen.

Was ich bauen würde, nach Rolle

SoC- und TV-Chip-Teams. Behalten Sie festverdrahtete Decoder für jeden Standard, den Sie abspielen müssen, darunter jetzt AV2 und VVC, wo der Rundfunk es verlangt. Dimensionieren Sie die NPU für Nachverarbeitung in Displayauflösung und -bildrate, und legen Sie den Pfad vom Decoder über die NPU zum Display so aus, dass die Bildverbesserung keinen zusätzlichen Umweg über den DRAM erzeugt. Unterstützen Sie SEI für neuronale Postfilter. Sehen Sie deterministische Ganzzahlmodi in der NPU vor. Behandeln Sie H.267 als Designstart 2029 oder später, und halten Sie etwas Programmierbarkeit für seine neuronalen Werkzeuge vor.

IP-Anbieter. Verkaufen Sie jetzt AV2- und VVC-Decoder, dazu die Bausteine, die neuronale Verarbeitung um sie herum braucht: Postfilter-Beschleuniger, Framebuffer-Kompression und Konformitätswerkzeuge.

Videoplattformen. Neuronale Vorverarbeitung am Encoder zahlt sich zuerst aus, weil sie einmal pro Titel läuft und nichts von den Geräten verlangt. Nutzen Sie clientseitige Super-Resolution, wo die Geräte es zulassen. Behalten Sie neuronale End-to-End-Codecs geschlossenen Systemen vor, die Sie kontrollieren.

KAPITEL 09

STANDARDS UND PATENTE

09 Standards und Patente: wo neuronale Werkzeuge lizenzierbar werden

Die Standardisierungs-Pipeline

Standard	Gremium	Neuronaler Inhalt	Stand November 2026
H.267 (Arbeitsname; jenseits von VVC)	JVET (ITU-T und ISO/IEC)	Offen. Klassische Werkzeuge (ECM) und neuronale Werkzeuge (In-Loop-Filter, Intra-Prädiktion, Super-Resolution) sind beide Kandidaten	Call for Proposals Juli 2026; Auswertung der Einreichungen Januar 2027; erstes Testmodell zweite Jahreshälfte 2027; Version 1 Ende 2029
VVC und HEVC mit VSEI	JVET	Ein neuronaler Postfilter kann in SEI-Metadaten beschrieben werden; ihn anzuwenden ist für den Decoder optional	In Kraft; Nachrichten in den Ausgaben vom Oktober 2025 erweitert
AV2	AOMedia	Klassische Werkzeuge, lizenzgebührenfrei	Version 1.0.0 am 28. Mai 2026; Decoder-IP ab Juni 2026 lizenziert; YouTube-Webstreams für Anfang 2027 geplant
JPEG AI	JPEG (ISO/IEC und ITU-T)	Gelernter End-to-End-Bildcodec für menschliches und maschinelles Sehen	Teil 1 veröffentlicht 2025; eine zweite Version in Diskussion; keine standardmäßige Unterstützung in Produkten gefunden
IEEE 1857.11	IEEE	Bildcodierung mit neuronalen Netzen, ausgelegt für NPUs	Veröffentlicht im Dezember 2024
Video Coding for Machines	MPEG	Codierung für maschinelle Analyse statt für Betrachter	Internationaler Standardentwurf Anfang 2026; endgültige Fassung für Ende 2026 geplant
Opus-ML-Erweiterungen	IETF	Optionale Deep Redundancy, decoderseitige Verbesserung, Bandbreitenerweiterung	In libopus 1.5 und 1.6; Deep Redundancy in der Standardisierung

Zwei Dinge fallen auf. Erstens gelangen neuronale Inhalte über die Ränder in die Standards, bevor sie den Kern erreichen: als optionale Postfilter, optionale Verbesserung und Bildcodierung, die alle den Hauptdecodierpfad unverändert lassen. Zweitens ist der eine Standard, der neuronale Werkzeuge in den Decodierpfad jedes Bildschirms bringen könnte, H.267, drei Jahre von einer ersten Version entfernt.

Wer bei den Patenten in Position ist

Standardwerkzeuge werden zu lizenzierbaren Patenten, deshalb ist der H.267-Prozess auch ein Patentrennen, und die neuronalen Werkzeuge gehören dazu. Die Patentlandschaft von GreyB aus dem Jahr 2025 über 3.916 Patente jenseits von VVC sieht ByteDance vorn, mit 645 Patentfamilien und 251 JVET-Beiträgen, und verweist auf Partnerschaften zwischen Hochschulen und Industrie, etwa die von Xiaomi und der Xidian University mit mehr als 20 Familien. MLex berichtete im März 2026, dass Tencent, Alibaba und ByteDance zu aktiven Teilnehmern an Pa-

tentpools für Videocodecs werden. Im Mai 2026 kaufte Sauvegarder Investment Management von Alibaba 126 Videocodec-Patente, die AV1, AV2 und Video Coding for Machines abdecken. Mit dem Kauf von Deep Render im Oktober 2025 hat InterDigital dessen Patente zur KI-Videocodierung in sein Portfolio aufgenommen.

Das Geld pro Gerät lässt sich in den Gebührentabellen der Pools leicht ablesen. Die Tarife von Access Advance für VVC (Stand 1. Juli 2026, für vertragstreue Lizenznehmer, die das Markenzeichen nutzen) betragen 0,50 US-Dollar pro Mobiltelefon, Tablet oder Laptop in Märkten der Region 1 wie den USA, der EU und Japan und 0,25 US-Dollar in Region 2; für 4K-Fernseher je nach Preis 1,00 bis 1,50 US-Dollar in Region 1; dazu gilt eine jährliche Obergrenze von 75 Millionen US-Dollar pro Unternehmen. AV1 und AV2 sind seitens der AOMedia-Mitglieder lizenzgebührenfrei. Was auch immer H.267 enthält, seine Werkzeuge werden auf die gleiche Weise bepreist, und deshalb bringen sich Unternehmen mit Patenten zur neuronalen Codierung schon jetzt bei JVET ein Analyse.

KAPITEL 10

10 WER VERDIENT
Wer verdient, und wo ein Unternehmen ansetzen kann

Das Geld im Jahr 2026

ABB. 09 **Wer 2026 an neuronaler Codierung verdient**

Das Geld folgt den Standards und den Plattformen. Ich habe kein eigenständiges Unternehmen gefunden, das von der Lizenzierung eines neuronalen End-to-End-Videocodecs lebt.

VERDIENEN JETZT

Plattformen

Behalten die Ersparnis

Neuronaler Downscaler von Netflix; Super-Resolution bei YouTube und Meta; Microsoft hat MLVC unter MIT-Lizenz offengelegt

VERDIENEN JETZT

Geräte- und Chiphersteller

Verkaufen Differenzierung

Kamera-ISPs; TV-Upscaling-Prozessoren; Tencents hauseigene Canghai-Codec-Chips

VERDIENEN JETZT

IP-Anbieter

Lizenzieren jeden neuen Standard

AV2-Decoder-Kunden bei Chips&Media, Allegro DVT und VeriSilicon binnen drei Wochen nach der Spezifikation

VERDIENEN JETZT

GPU-Hersteller

Schaffen Plattformbindung

DLSS 4 in über 250 Spielen und Apps; SDK für neuronale Texturen in der Beta

PRO GERÄT

Patentinhaber und Pools

Kassieren pro Gerät

VVC: 0,50 $ pro Smartphone, 1,00–1,50 $ pro 4K-TV; H.267-Werkzeuge nach 2029

EXIT DURCH ÜBERNAHME

Start-ups für neuronale Codecs

Übernommen

WaveOne → Apple (2023); iSIZE → Sony (2023); Deep Render → InterDigital (2025)

Quelle: Im Bericht zitierte Angaben von Unternehmen und Patentpools; Stand 12. November 2026. Pool-Tarife: Access Advance, VVC, Region 1, vertragstreu mit Markennutzung, Stand 1. Juli 2026.

Plattformen behalten die Einsparungen. Netflix, YouTube, Meta, Microsoft und Tencent entwickeln neuronale Verarbeitung im eigenen Haus und behalten den Nutzen in Form geringerer Bandbreite, geringeren Speicherbedarfs und besserer Qualität pro Bit. Wenn es der Verbreitung hilft, verschenken sie den Code: Microsoft hat MLVC unter der MIT-Lizenz veröffentlicht.

Geräte- und Chiphersteller verkaufen Differenzierung. Kamera-ISPs und TV-Bildprozessoren sind die Stellen, an denen neuronale Netze jeden Nutzer erreichen, und sie werden als Features verkauft: Qualcomms Cognitive

ISP, Samsungs Upscaling-Prozessoren.

IP-Anbieter verkaufen Decoder für jeden neuen Standard, dazu NPUs. Binnen drei Wochen nach der AV2-Spezifikation meldeten Chips&Media, Allegro DVT und VeriSilicon allesamt Kunden für AV2-Decoder. Dieselben Anbieter verkaufen jetzt neuronale Verarbeitungsblöcke neben den Codecs.

GPU-Hersteller bauen Plattformen. DLSS und neuronale Texturkompression machen NVIDIAs Tensor-Hardware zu einem Teil der Spieleentwicklung.

Patentinhaber kassieren pro Gerät, über Pools für VVC und HEVC und nach 2029 für H.267.

Start-ups für neuronale Codecs wurden übernommen. WaveOne ging 2023 an Apple, iSIZE 2023 an Sony Interactive Entertainment und Deep Render 2025 an InterDigital, alle zu nicht veröffentlichten Bedingungen. In jedem Fall kaufte eine Plattform oder ein Lizenzgeber das Team und die Patente. Ich habe 2026 kein eigenständiges Unternehmen gefunden, das von der Lizenzierung eines neuronalen End-to-End-Videocodecs lebt.

Wo ein Unternehmen ansetzen kann

Geschäftsfeld	Wer kauft	Warum jetzt	Hauptrisiko
Neuronale Vorverarbeitung am Encoder (Downscaling, Entrauschen, titelspezifisches Tuning)	Streaming- und Nutzervideo-Plattformen, Sender	Keine Geräteänderung nötig; bei Netflix erprobt	Große Plattformen bauen es selbst
IP und Tuning für displayseitige Bildverbesserung (Super-Resolution, Zwischenbildberechnung, Bildqualität)	Chiphersteller für TVs, Set-Top-Boxen und Smartphones	NPUs stecken schon in den Geräten; sichtbare Differenzierung	Kommoditisierung; Chiphersteller entwickeln selbst
Neuronale Codecs für geschlossene Systeme (Konferenzen, Fernbetrieb, private Verbindungen)	Anbieter für Unternehmenskommunikation, Industrie und Telepräsenz	MLVC beweist die Machbarkeit, und sein Code ist offen	Sie müssen beide Enden kontrollieren; NPU-Fragmentierung
Codec plus NPU als Silizium oder IP (AV2, VVC und Nachverarbeitung)	SoC-Hersteller	Der AV2-Zyklus beginnt	Starke etablierte Anbieter
Werkzeuge für Evaluierung, Konformität und deterministische Inferenz	Alle, die neuronale Werkzeuge ausliefern	Bitgenauigkeit ist eine harte Anforderung	Kleiner Markt, außer im Paket
Standardbeiträge zu H.267	Lizenzgeber, große Plattformen	Die Einreichungen werden im Januar 2027 ausgewertet	Langwierig, teuer, ungewiss

Für die meisten Unternehmen ist 2027 in den ersten beiden Zeilen Umsatz zu holen. Die dritte ist für ein fokussiertes Produktunternehmen die interessanteste. Die letzte braucht Geduld und eine Lizenzstrategie.

KAPITEL 11

BEHAUPTUNGEN UND BELEGE

11 Behauptungen, die Sie hören werden, und was die Belege sagen

Behauptung	Was die Belege sagen
„Neuronale Codecs schlagen VVC bereits.“	Im Labor, ja. DCVC-RT spart im Testsetup seines Papers bei 1080p 21 % gegenüber dem VVC-Referenzencoder; JPEG AI spart 12,5–27,9 % gegenüber VVC Intra. Der Decoder kostet etwa 175 Tausend MACs pro Pixel oder die Energie einer NPU, gegenüber einem Bruchteil eines Nanojoule in festverdrahteter Hardware.
„KI-Codecs werden H.26x und AV1 in wenigen Jahren ersetzen.“	Nicht bei der Wiedergabe. Der erste Standard, der neuronale Werkzeuge tragen kann, zielt auf Ende 2029, Chips folgen um 2031 und der breite Einsatz Mitte der 2030er-Jahre Schätzung. Der Codec, der jetzt kommt, AV2, ist klassisch.
„MLVC spart 88 % Bandbreite.“	Gegenüber H.264 bei 360p, auf Microsofts eigenem Testdatensatz. Gegenüber Hardware-HEVC sind es 75,5 % bei 360p und 65,4 % bei 540p, in Anrufen, bei denen Microsoft beide Enden kontrolliert.
„Neuronale Texturen sparen 85–96 % Grafikspeicher.“	In Demonstrationsszenen. Das ausgelieferte Spiel gewann etwa 30 % effektive Kompression und nutzte neuronale Texturen wegen der Laufzeitkosten nur für ausgewählte Assets.
„JPEG AI ist ein ISO-Standard, also werden Kameras und Browser ihn übernehmen.“	Ich habe keine standardmäßige Unterstützung in Produkten gefunden. Das Format, das das Web 2026 übernommen hat, JPEG XL in Chrome 155, ist klassisch.
„Neuronale Decodierung ist für Smartphones zu teuer.“	Teilweise richtig. MLVC decodiert 540p30 mit einer halben NPU; 4K60 mit einem Modell der DCVC-RT-Klasse bräuchte rund 175 TOPS Dauerrechenleistung Schätzung.
„Satelliten- und Drohnenverbindungen sind der große Markt für neuronales Video.“	Technisch die beste Eignung, aber klein und auf beiden Seiten des Pazifiks zunehmend Exportkontrollen unterworfen.
„Große Plattformen werden einen neuronalen Codec von einem Start-up lizenzieren.“	Sie haben ihren eigenen gebaut (Netflix, Microsoft, Tencent) oder das Start-up gekauft (Apple, Sony, InterDigital).

Ein einfacher Test entzaubert die meisten Pitches: Fragen Sie nach Referenz, Metrik, Auflösung und Decodieraufwand. Eine Behauptung, die nicht alle vier nennen kann, ist ein Forschungsergebnis, kein Produkt.

KAPITEL 12

12

WEGMARKEN

Worauf zu achten ist, 2027 bis 2030

Wann	Wegmarke	Warum sie wichtig ist
Januar 2027	JVET wertet die H.267-Einreichungen in Brisbane aus	Wie viele der besten Einreichungen auf neuronale Werkzeuge setzen, und zu welchen Kosten pro Pixel
Anfang 2027	Erste AV2-Webstreams bei YouTube; AV2-Decoder in angekündigten SoCs	Der nächste klassische Zyklus legt die Gerätebudgets bis etwa 2030 fest
Zweite Jahreshälfte 2027	Erstes Testmodell von H.267	Ob neuronale Werkzeuge im Kerndesign stecken oder nur in der Exploration
2027	MLVC über Peer-to-Peer-Anrufe in Teams hinaus; andere Anbieter übernehmen den offenen Code	Die erste Bewährungsprobe eines offenen neuronalen Codecs außerhalb seines Eigentümers
2027–2028	Ein erstes Produkt, das JPEG AI standardmäßig schreibt oder liest	Das erste neuronale Format, das über Unternehmensgrenzen hinweg genutzt wird
2027–2028	DirectX Linear Algebra in ausgelieferten Treibern; neuronale Texturen in mehr Spielen und auf Mobilgeräten	Herstellerübergreifende neuronale Decodierung in der Grafik
Laufend	NPU-Budgets in Smartphones und PCs	Bei rund 100 TOPS Dauerleistung würde ein 1080p30-Decoder der DCVC-RT-Klasse etwa ein Fünftel der NPU belegen statt der Hälfte Analyse
Ende 2029	H.267 Version 1	Der Beginn des Chipzyklus für neuronale Werkzeuge in der Wiedergabe im Massenmarkt

KAPITEL 13

13 Offene Fragen

Mehrere Fragen in diesem Bericht haben noch keine öffentliche Antwort, und gerade sie finde ich technisch am interessantesten.

Exaktheit im großen Maßstab. Wie nah kann ein neuronaler Decoder an bitgenaue, reine Ganzzahlarithmetik über NPUs verschiedener Hersteller herankommen, bevor er den Großteil seines Kompressionsvorteils aufgibt? Der Tausch von sechs Prozentpunkten bei MLVC ist ein Datenpunkt; ein Standard braucht eine allgemeine Antwort.

Speicher vor MACs. Was kostet neuronale Nachverarbeitung an DRAM-Verkehr pro Pixel, wenn sie im Durchlauf zwischen einem Hardware-Decoder und dem Display läuft, und wie klein lässt sich ihr Working Set machen?

Der Preis eines neuronalen Werkzeugs in H.267. Ab wie vielen MACs und Nanojoule pro Pixel bezahlt ein neuronaler In-Loop-Filter in einem 4K60-Decoder nicht mehr die Bits, die er spart?

Geschlossene Systeme jenseits der Telefonie. Welche anderen geschlossenen Systeme, von Cloud-Fotospeichern und Game-Streaming bis zu Robotern und privaten Zuspielstrecken, können neuronale End-to-End-Codecs übernehmen, bevor H.267 kommt?

Decoder, die generieren. Wenn ein Decoder Details erzeugt, statt sie zu rekonstruieren, wie es Video-VAEs bereits tun, wie sollte Wiedergabetreue gemessen und spezifiziert werden?

Das sind Felder, die wir in den kommenden Jahren gemeinsam erkunden können, im technischen Austausch und in technischer Zusammenarbeit: i@tanshuai.com.

Grenzen dieses Berichts

Dies ist eine Sekundärrecherche auf Basis öffentlicher Dokumente bis zum 12. November 2026. Für diese Ausgabe habe ich keine Primärinterviews und keine unabhängigen Messungen durchgeführt. Mehrere Schlüsselzahlen sind Angaben von Anbietern oder Autoren auf ihren eigenen Testdatensätzen, etwa die Bitrateneinsparungen von MLVC und die Kompressionsraten neuronaler Texturen; sie sind als solche gekennzeichnet. Meine Schätzungen des Rechenaufwands für DCVC-RT addieren die Werte pro Modul, die seine Autoren veröffentlichen, und ignorieren Auslastungsverluste, sodass die reale NPU-Last höher ausfallen wird. Der Energievergleich stützt sich auf wenige veröffentlichte Designs in unterschiedlichen Fertigungsprozessen; Produktionswerte für VVC- und AV2-Decoder sind proprietär. Die Zeitpläne der Standards sind die eigenen Ziele der Gremien. Die H.267-Daten nach 2029 sind meine Schätzungen auf Basis der Geschichte von AV1 und VVC.



Über den Autor

Tan Shuai ist Technologe und Gründer mit mehr als zwanzig Jahren Arbeit an Chips, Hardware und Systemsoftware und 20 erteilten Patenten. Er hat viele Jahre Chipentwicklung und Management im Fabless-Bereich hinter sich, darunter den Full-Custom-Entwurf von ASIC-Datenpfaden und Multimedia-Kamerachip-Projekte wie Beschleuniger für Videocodecs; daraus ergibt sich der Blickwinkel dieses Berichts: Energie pro Pixel, Speicherverkehr und was ein Standard einen Decoder in Silizium festschreiben lässt. Seine Arbeit führte ihn ins Silicon Valley, nach New York, Chicago und Taiwan, in die Greater Bay Area (Shenzhen und Hongkong) sowie nach Vietnam und Malaysia. Heute arbeitet er als KI-Forscher und -Entwickler.

Technische Diskussion: i@tanshuai.com · tanshuai.com

Hinweise und Haftungsausschluss

Dieser Bericht ist eine unabhängige Sekundärrecherche auf Basis öffentlicher Dokumente, die bis zum 12. November 2026 verfügbar waren. Er dient ausschließlich der Information und ist keine Anlage-, Rechts- oder technische Beratung. Produkte, Zeitpläne von Standards, Tarife von Patentpools und Exportregeln ändern sich; prüfen Sie alles, worauf Sie sich verlassen, an der Primärquelle. Mit Anbieterangabe, Schätzung oder Analyse gekennzeichnete Zahlen sind entsprechend zu lesen. Firmen- und Produktnamen gehören ihren Inhabern und dienen nur der Identifikation; kein Unternehmen hat diesen Bericht geprüft oder gesponsert.

© 2026 Tan Shuai. Sie dürfen diesen Bericht unverändert und mit Quellenangabe weitergeben.

Quellen

Alle abgerufen zwischen dem 3. und dem 12. November 2026. Nach Kapiteln gruppiert.

Datenverkehr, Geräte und Abgrenzung

- Ericsson, *Mobility Report*, Juni 2026 (Video etwa 75 % des mobilen Datenverkehrs Ende 2025): https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2026 ; Zusammenfassung: https://telecomlead.com/5g/ericsson-mobility-report-2026-global-mobile-data-traffic-to-reach-515-eb-by-2031-as-india-leads-smartphone-data-consumption-126290
- IDC, weltweite Smartphone-Auslieferungen, Q4 und Gesamtjahr 2025 (1,26 Milliarden), 13. Januar 2026: https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce, Prognose der TV-Auslieferungen 2025 (195,71 Millionen): https://iconnect007.com/article/146406/early-shipments-by-tv-brands-push-global-1h25-volume-up-2-but-raise-concerns-for-secondhalf-peak-season/146403/ein ; Omdia, TV-Prognose 2025 (208,7 Millionen): https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

Wiedergabe und Streaming

- Netflix Technology Blog, „AV1 — Now Powering 30% of Netflix Streaming“, Dezember 2025: https://netflixtechblog.com/av1-now-powering-30-of-netflix-streaming-02f592242d80
- Netflix Technology Blog, „For your eyes only: improving Netflix video quality with neural networks“, November 2022: https://netflixtechblog.com/for-your-eyes-only-improving-netflix-video-quality-with-neural-networks-5b8d032da09c
- AOMedia, AV2 Bitstream and Decoding Process Specification v1.0.0, 28. Mai 2026: https://av2.aomedia.org/v1.0.0/index.html ; „AV2: What the Final Specification Means for Implementers“: https://aomedia.org/blog%20posts/AV2-What-the-Final-Specification-Means-for-Implementers/
- TV Technology, „YouTube to Commence AV2 Streaming in 2027“ (AV1-Anteil am YouTube-Katalog): https://www.tvtechnology.com/platform/streaming/youtube-to-commence-av2-streaming-in-2027
- Google Research, „The in-loop filtering pipeline in AV2“, 2026: https://research.google/pubs/the-in-loop-filtering-pipeline-in-av2/
- Fraunhofer IIS, „The Future of Broadcast in Brazil“ (kommerzieller Betrieb von DTV+, Juni 2026): https://mpegh.com/brazil/ ; CartaCapital zu Empfängern in Fernsehern: https://www.cartacapital.com.br/?p=628337 ; Olhar Digital zu VVC in TV 3.0: https://olhardigital.com.br/2026/05/01/pro/em-breve-a-tv-aberta-vai-mudar-no-brasil/
- 9to5Google, „YouTube adds AI-upscaled 'Super Resolution' for low-quality videos“, 29. Oktober 2025: https://9to5google.com/2025/10/29/youtube-adds-ai-upscaled-super-resolution-for-low-quality-videos/
- Meta, „How Meta Deployed Super Resolution at Scale to Transform Video Quality“, @Scale, 16. Oktober 2025: https://atscaleconference.com/how-meta-deployed-super-resolution-at-scale-to-transform-video-quality/
- Samsung, 8K-TV-Highlights (NQ8 AI Gen3, 768 neuronale Netze): https://www.samsung.com/us/tvs/8k-tv/highlights/

Fotografie und Bilder

- Qualcomm, „Meet the Snapdragon 8 Gen 2 Super Cameras“ (Cognitive ISP): https://www.qualcomm.com/news/onq/2023/07/meet-the-snapdragon-8-gen-2-super-cameras
- Chips&Media und Visionary.ai, vollständig KI-basierter ISP, 5. Januar 2026: https://www.edge-ai-vision.com/2026/01/chipsmedia-and-visionary-ai-unveil-the-worlds-first-ai-based-full-image-signal-processor-redefining-the-future-of-image-quality/ ; Interview bei PetaPixel: https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- JPEG-Komitee, Pressemitteilung zum 103. Treffen (Gewinne, Komplexität und Smartphone-Implementierungen von JPEG AI): https://jpeg.org/items/20240507_press.html
- G. J. Sullivan, Einführung in ISO/IEC JTC 1/SC 29, April 2026 (Status von JPEG AI; Zahlen zu ECM und NNVC): https://www.mpeg.org/introduction-to-iso-iec-jtc-1-sc-29-coding-of-audio-picture-multimedia-and-hypermedia-information-the-world-of-jpeg-and-mpeg/
- Arbeitsgruppe IEEE 1857.11 (Bildcodierung auf Basis neuronaler Netze): https://sagroups.ieee.org/fvc
- Chrome 155 und JPEG XL, 6. Oktober 2026: https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

Telefonie und Konferenzen

- Microsoft, „Satin: Microsoft's latest AI-powered audio codec for real-time communications“, 17. Februar 2021: https://techcommunity.microsoft.com/t5/microsoft-teams-blog/satin-microsoft-s-latest-ai-powered-audio-codec-for-real-time/ba-p/2141382
- Microsoft, „Announcing the Open-Source Release of ML Video Codec (MLVC)“, 24. Juli 2026: https://techcommunity.microsoft.com/blog/linuxandopensourceblog/announcing-the-open-source-release-of-ml-video-codec-mlvc/4539875 ; Code: https://github.com/microsoft/mlvc
- Streaming Learning Center, „Microsoft MLVC: Interoperability, Power Realities, and the Neural Compression Patent Landscape“: https://streaminglearningcenter.com/articles/microsoft_mlvc.html ; Analyse von iReplay: https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/

- Release Notes zu Opus 1.5: https://opus-codec.org/demo/opus-1.5 ; Opus 1.6: https://opus-codec.org/demo/opus-1.6
- Meta Engineering, „MLow: Meta's low bitrate audio codec“, 13. Juni 2024: https://engineering.fb.com/2024/06/13/web/mlow-metas-low-bitrate-audio-codec/
- Microsoft, Copilot+ PCs (NPU mit 40+ TOPS): https://www.microsoft.com/en-us/windows/copilot-plus-pcs ; Apple, Ankündigung des M4 (Neural Engine mit 38 TOPS): https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/

Generative KI

- Kyutai, Model Card von Mimi (12,5 Hz, 1,1 kbps): https://huggingface.co/kyutai/mimi
- Lightricks, „LTX-Video: Realtime Video Latent Diffusion“ (Video-VAE mit 1:192): https://arxiv.org/abs/2501.00103
- Pressemitteilung zum 153. MPEG-Treffen (Video Coding for Machines im DIS-Stadium): https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

Spiele und Grafik

- NVIDIA, Verbreitung von DLSS 4: Start und Computex 2025 (75 und 125+): https://www.nvidia.com/en-us/geforce/news/125-dlss-4-multi-frame-gen-games-more-announced-computex-2025/ ; Gamescom 2025 (175+): https://www.nvidia.com/en-us/geforce/news/gamescom-2025-dlss-4-full-ray-tracing-game-announcements/ ; CES 2026 (250+, DLSS 4.5): https://www.nvidia.com/en-us/geforce/news/dlss-4-rtx-path-tracing-game-announcements-ces-2026/
- Ubisoft La Forge, „Shipping Neural Texture Compression in Assassin's Creed Mirage“, Februar 2026: https://www.ubisoft.com/en-us/studio/laforge/news/415slmB3ZGzv8d2LhO9QT/shipping-neural-texture-compression-in-assassins-creed-mirage
- TweakTown zur Demo neuronaler Texturen von NVIDIA auf der GTC 2026 (6,5 GB auf 970 MB): https://www.tweaktown.com/news/110886/nvidias-neural-texture-compression-cuts-vram-usage-from-6-5gb-down-to-970mb/index.html
- Microsoft DirectX Blog, Shader Model 6.9 Retail, 26. Februar 2026: https://devblogs.microsoft.com/directx/shader-model-6-9-retail-and-more/ ; zu Cooperative Vector: https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ; Pläne für die GDC 2026: https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

Spezialverbindungen

- Global Trade Alert zu Chinas Bekanntmachung Nr. 49 von 2026: https://globaltradealert.org/intervention/161396-china-government-revises-export-controls-on-drone-related-items ; DroneXL: https://dronexl.co/2026/10/19/china-drone-export-controls-motors-props-november-18/ ; Army Recognition (Schwellenwerte für unbemannte Bodenplattformen): https://www.armyrecognition.com/news/army-news/2026/china-tightens-drone-export-controls-raising-supply-chain-risks-for-the-us-europe-and-ukraine
- DroneLife, FCC nimmt im Ausland hergestellte Drohnen und Komponenten in die Covered List auf, 22. Dezember 2025: https://dronelife.com/2025/12/22/fcc-adds-foreign-made-drones-and-components-to-covered-list-citing-national-security-risks/

Silizium

- M. Tikekar et al., „A 249-Mpixel/s HEVC Video-Decoder Chip for 4K Ultra-HD Applications“, IEEE JSSC 2014: https://dspace.mit.edu/handle/1721.1/93876 ; Chipseite des MIT (78 mW bei 4K30): https://eecsg.mit.edu/chips/hevc-video-decoder-for-4k-ultra-hd-applications/
- S. Zhang et al., NVCA, DATE 2024: https://arxiv.org/abs/2312.10716
- Electronic Design, „Why Traditional VVC Will Outlive the Neural-Codec Hype on Silicon“, 2026: https://www.electronicdesign.com/home/article/55414955/why-traditional-vvc-will-outlive-the-neural-codec-hype-on-silicon
- Z. Jia et al., „Towards Practical Real-Time Neural Video Compression“ (DCVC-RT), CVPR 2025: https://openaccess.thecvf.com/content/CVPR2025/html/Jia_Towards_Practical_Real-Time_Neural_Video_Compression_CVPR_2025_paper.html ; Zusatzmaterial (Komplexität pro Modul): https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVC und die Ganzzahl-Inferenzbibliothek SADL: https://arxiv.org/abs/2309.05846
- Tencent Canghai V2, 27. Mai 2026 (Shanghai Securities News via 10jqka): https://stock.10jqka.com.cn/20260527/c677016020.shtml ; Zusammenfassung der Funktionen: https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&Media, AV2-Decoder-IP, 12. Mai 2026: https://www.design-reuse.com/news/202530501-chips-media-completes-development-of-next-gen-av2-hw-decoder-ip/ ; AV2-Lizenz, 17. Juni 2026: https://www.edge-ai-vision.com/?p=58270
- Allegro DVT, AV2-Decoder-IP, 16. Juni 2026: https://www.edge-ai-vision.com/2026/06/allegro-dvts-pulsar-decoder-ip-adds-support-for-av2-video-codec/
- VeriSilicon, VC9800D mit AV2, 10. Juni 2026: https://www.businesswirechina.com/en/news/63176.html

Standards und Patente

- Pressemitteilung zum 155. MPEG-Treffen (gemeinsamer Call for Proposals jenseits von VVC), 5. August 2026: https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- Pressemitteilung zum 152. MPEG-Treffen (Ergebnisse des Call for Evidence; Ausgaben von VVC und VSEI), 30. Oktober 2025: https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ; MPEG-Kolumne von ACM SIGMM: https://records.sigmm.org/2025/12/09/mpeg-column-152nd-mpeg-meeting/

- Streaming Media, „AI Helps InterDigital Reach Beyond VVC in Race to Develop Next-Gen Codec“: https://www.streamingmedia.com/Articles/News/Online-Video-News/AI-Helps-InterDigital-Reach-Beyond-VVC-in-Race-to-Develop-Next-Gen-Codec-171971.aspx
- GreyB, Patentlandschaft jenseits von VVC (H.267): https://greyb.com/blog/exclusive-video-codec-landscape/ ; zur Lizenzierung von AV2 und zum Verkauf des Alibaba-Portfolios: https://greyb.com/blog/av2-licensing/
- MLex, „Fragmented codec licensing exposes need for clearer rules for Chinese tech majors“, 27. März 2026: https://www.mlex.com/mlex/articles/2458402/fragmented-codec-licensing-exposes-need-for-clearer-rules-for-chinese-tech-majors
- Access Advance, Lizenzgebühren für MCBA und VVC Advance, Stand 1. Juli 2026: https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

Geld und Unternehmen

- InterDigital übernimmt Deep Render, 30. Oktober 2025: https://www.advanced-television.com/2025/10/30/interdigital-acquires-deep-render
- TechCrunch, „Apple acquired a startup using AI to compress videos“, 27. März 2023: https://techcrunch.com/2023/03/27/apple-acquired-a-startup-using-ai-to-compress-videos
- Neowin, „Sony acquires AI video startup iSIZE“, November 2023: https://www.neowin.net/news/sony-acquires-ai-video-startup-isize-to-enhance-playstation-streaming/

GELERNTE KOMPRESSION · WIEDERGABE × SILIZIUM

Neuronale Codierung 2026

Im Einsatz rund um den Codec, nicht an seiner Stelle

Jede wichtige Zahl in diesem Bericht trägt Quelle und Datum. Wo der Autor eine Zahl ableitet, ist sie als Schätzung gekennzeichnet; wo nur der Verkäufer sie angibt, als Anbieterangabe.

Tan Shuai

Technologe und Gründer · KI-Forscher und -Entwickler

tanshuai.com · i@tanshuai.com

ÖFFENTLICHE AUSGABE · Datenstand: 12. November 2026